

**KLASIFIKASI PENGIDAP KANKER PAYUDARA
MENGUNAKAN METODE *VOTING BASED EXTREME
LEARNING MACHINE* (V-ELM)**

SKRIPSI

Untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun oleh:
Dheby Tata Artha
NIM: 155150207111068



PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BRAWIJAYA
MALANG
2018

PENGESAHAN

KLASIFIKASI PENGIDAP KANKER PAYUDARA MENGGUNAKAN METODE VOTING
BASED EXTREME LEARNING MACHINE (V-ELM)

SKRIPSI

Untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

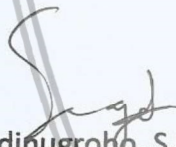
Disusun Oleh :
Dheby Tata Artha
NIM: 155150207111068

Skrisi ini telah diuji dan dinyatakan lulus pada
28 Desember 2018

Telah diperiksa dan disetujui oleh:

Pembimbing I

Pembimbing II



Sigit Adinugroho, S.Kom., M.Sc.
NIK: 2016078807011001


Putra Pandu Adikara, S. Kom., M.Kom.
NIP: 19850725-200812 1 002

Mengetahui

Ketua Jurusan Teknik Informatika



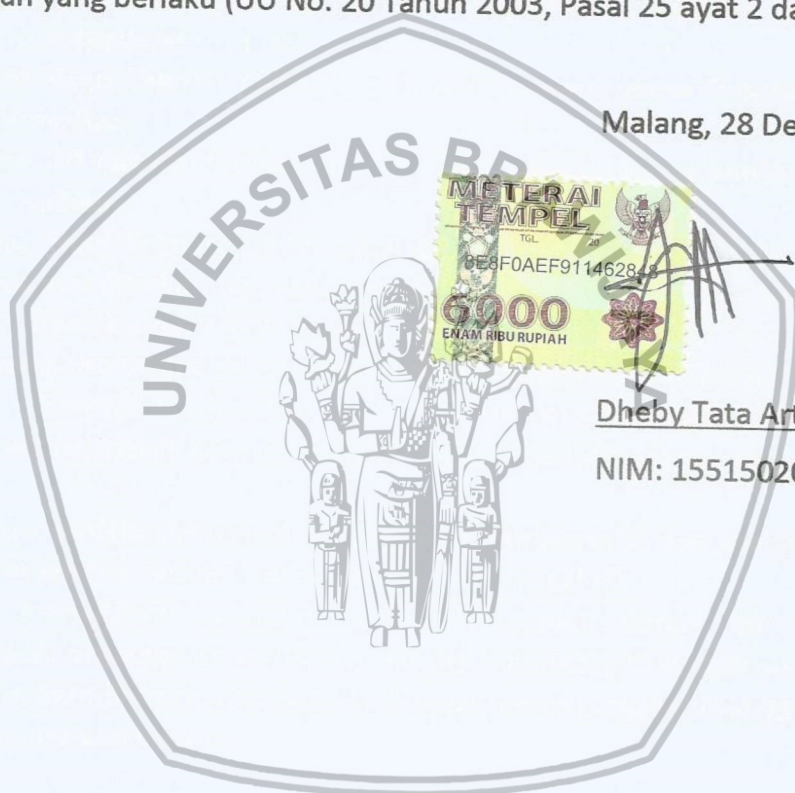

Tri Astoto Kurniawan, S.T., M.T., Ph.D.
NIP: 19710518 200312 1 001

PERNYATAAN ORISINALITAS

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar referensi.

Apabila ternyata didalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur plagiasi, saya bersedia skripsi ini digugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70).

Malang, 28 Desember 2018



Dheby Tata Artha

NIM: 155150207111068

PRAKATA

Puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat, taufik dan hidayah-Nya sehingga laporan skripsi yang berjudul “Klasifikasi Pengidap Kaker Payudara Menggunakan Metode *Voting Based Extreme Learning Machine* (V-ELM)” ini dapat terselesaikan.

Penulis menyadari bahwa skripsi ini tidak akan berhasil tanpa bantuan dan peran serta beberapa pihak. Oleh karena itu, penulis ingin menyampaikan rasa hormat dan terima kasih kepada:

1. Bapak Sigit Adinugroho, S.Kom., M.Sc., dan Bapak Putra Pandu Adikara, S.Kom., M.Kom. selaku Pembimbing skripsi yang telah dengan sabar dan pengertian membimbing serta mengarahkan penulis sehingga dapat menyelesaikan skripsi ini,
2. Bapak Agus Wahyu Widodo, S.T., M.Cs. selaku ketua Program Studi Teknik Informatika,
3. Bapak Tri Astoto Kurniawan, S.T., M.T., Ph.D. selaku ketua Jurusan Teknik Informatika,
4. Bapak Fajar Pradana, S.ST., M.Eng. selaku dosen Penasihat Akademik yang selalu memberikan nasihat kepada penulis selama menempuh masa studi,
5. Ayah dan Ibu dari penulis atas segala doa, dukungan dan semangat yang diberikan sehingga penulis dapat mengerjakan laporan skripsi ini,
6. Seluruh civitas akademika Teknik Informatika Universitas Brawijaya yang telah banyak memberi bantuan dan dukungan selama penyelesaian laporan skripsi ini,
7. Sukma Fardhia Anggraini yang selalu memberikan doa, semangat, motivasi dan selalu menemani dalam pengerjaan skripsi ini.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak kekurangan, sehingga saran dan kritik yang membangun sangat penulis harapkan. Akhir kata penulis berharap skripsi ini dapat membawa manfaat bagi semua pihak yang membutuhkannya.

Malang, 28 Desember 2018

Penulis

tata.artha@gmail.com

ABSTRAK

Dheby Tata Artha. Klasifikasi Pengidap Kanker Payudara Menggunakan Metode *Voting Based Extreme Learning Machine* (V-ELM).

Pembimbing: Sigit Adinugroho, S.Kom., M.Sc. dan Putra Pandu Adikara, S.Kom., M.Kom.

Kanker payudara adalah tumor ganas yang berasal dari pertumbuhan abnormal sel-sel payudara. Setiap tahunnya, sebanyak 2,1 juta perempuan meninggal dunia akibat kanker payudara. Untuk mengurangi jumlah kematian akibat kanker payudara, dapat dilakukan upaya pencegahan dengan melakukan *screening*. Perkembangan teknologi medis dan teknologi informasi, di dalam dunia medis dapat digunakan peneliti di bidangnya untuk mengembangkan model deteksi dini, dari data konsultasi rutin dan analisis darah. Pada penelitian ini, dilakukan klasifikasi terhadap pengidap kanker payudara dengan metode *Voting Based Extreme Learning Machine* (V-ELM). Data yang digunakan adalah *Breast Cancer Coimbra Dataset* yang dipublikasikan pada *UCI Machine Learning* tahun 2018. Dalam data tersebut, terdapat 116 data dengan 9 fitur dan 2 kelas yaitu *Healthy Control* dan *Patient*. Data tersebut dilakukan normalisasi terlebih dahulu, kemudian dilakukan proses pelatihan V-ELM terhadap data latih. Setelah itu pengujian dilakukan dengan masukan berupa data uji serta hasil dari proses pelatihan. Rasio data latih dan data uji yang digunakan adalah 80:20. Penelitian ini melakukan pengujian terhadap beberapa parameter dan mendapatkan hasil optimal, diantaranya 20 *hidden* neuron, nilai *k* untuk V-ELM adalah sebesar 35 dan fungsi aktivasi dengan hasil optimal adalah fungsi *Sigmoid*. Dengan menggunakan ketiga parameter optimal tersebut didapatkan hasil rata-rata akurasi sebesar 89,56%, hasil rata-rata *sensitivity* sebesar 96,924% dan hasil rata-rata *specificity* sebesar 80%.

Kata kunci: klasifikasi, *Voting Based Extreme Learning Machine*, kanker payudara

ABSTRACT

Dheby Tata Artha. Breast Cancer Sufferers Classification Using Voting Based Extreme Learning Machine (V-ELM).

Supervisors: Sigit Adinugroho, S.Kom., M.Sc. and Putra Pandu Adikara, S.Kom., M.Kom.

Breast cancer is a malignant tumor that formed by the abnormal growth of breast cells. Every year, breast cancer causes about 2,1 million women to die. To reduce the number of deaths caused by breast cancer, screening can be chosen for prevention efforts. The development of medical technology and information technology, in the medical world, can be used by researchers in their fields to develop early detection models, from routine consultation data and blood analysis. In this study, breast cancer data will be classified using the Voting Based Extreme Learning Machine (V-ELM). This study using Coimbra Dataset Breast Cancer which published on UCI Machine Learning in 2018. It consists of 116 data, 9 features and 2 classes (Healthy Control and Patient). Firstly, the dataset would be normalized, then began the training process of V-ELM with data train. After that, began the testing process of V-ELM with input values from the training process and data test. The ratio between training data and testing data in this study is 80:20. This study tested several parameters and obtained optimal results, including 20 hidden neurons, the value of k for V-ELM is 35 and the activation function with optimal results is the Sigmoid function. By using those optimal parameters, gives average of accuracy is 89.56%, average of sensitivity is 96.924% and average of specificity is 80%.

Keywords: *classification, Voting Based Extreme Learning Machine, breast cancer*

DAFTAR ISI

PENGESAHAN	ii
PERNYATAAN ORISINALITAS	iii
PRAKATA.....	iv
ABSTRAK.....	v
ABSTRACT.....	vi
DAFTAR ISI	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xii
BAB 1 PENDAHULUAN.....	1
1.1 Latar belakang.....	1
1.2 Rumusan masalah.....	2
1.3 Tujuan	2
1.4 Manfaat.....	2
1.5 Batasan masalah	2
1.6 Sistematika Pembahasan.....	3
BAB 2 LANDASAN KEPUSTAKAAN	4
2.1 Kajian Pustaka	4
2.2 Payudara	5
2.2.1 Kanker Payudara	5
2.2.2 Faktor Kanker Payudara.....	5
2.3 Klasifikasi.....	6
2.4 Normalisasi Data	6
2.5 Jaringan Saraf Tiruan	7
2.6 <i>Extreme Learning Machine</i> (ELM).....	7
2.6.1 Arsitektur ELM.....	7
2.6.2 Fungsi Aktivasi.....	7
2.6.3 Proses <i>Training</i> ELM.....	8
2.6.4 Proses <i>Testing</i> ELM	9
2.7 <i>Voting Based Extreme Learning Machine</i> (V-ELM).....	9

2.7.1 Algoritme V-ELM	10
2.8 Nilai Evaluasi	10
BAB 3 METODOLOGI	12
3.1 Tipe penelitian	12
3.2 Strategi dan Rancangan Penelitian	12
3.2.1 Metode Secara Umum	12
3.2.2 Lokasi Penelitian.....	13
3.2.3 Metode Pengumpulan Data	13
3.2.4 Metode Analisis Data	13
3.2.5 Peralatan Pendukung	14
BAB 4 PERANCANGAN.....	15
4.1 Formulasi Permasalahan.....	15
4.2 Penyelesaian Masalah	15
4.2.1 Normalisasi Data	15
4.2.2 Proses Pelatihan V-ELM	17
4.2.3 Proses Pengujian V-ELM.....	19
4.3 Perhitungan Manualisasi	20
4.3.1 Perhitungan Manual Proses Pelatihan algoritme V-ELM.....	21
4.3.2 Perhitungan Proses Pengujian algoritme V-ELM	28
4.3.3 Perhitungan Proses Evaluasi algoritme V-ELM	32
4.4 Perancangan Pengujian	33
4.4.1 Pengujian Banyak <i>Hidden Neuron</i>	33
4.4.2 Pengujian Nilai <i>k</i>	33
4.4.3 Pengujian Fungsi Aktivasi	34
BAB 5 IMPLEMENTASI	35
5.1 Kode Sumber.....	35
5.1.1 Implementasi Normalisasi.....	35
5.1.2 Implementasi V-ELM	35
5.2 Antarmuka	39
BAB 6 PENGUJIAN DAN ANALISIS.....	40
6.1 Pengujian	40
6.2 Hasil dan Analisis Pengujian.....	40

6.2.1 Pengujian Banyak <i>Hidden Neuron</i>	40
6.2.2 Pengujian Fungsi Aktivasi	41
6.2.3 Pengujian Nilai <i>k</i>	43
6.2.4 Hasil Pengujian dan Analisis Global	44
BAB 7 PENUTUP	45
7.1 Kesimpulan	45
7.2 Saran	45
DAFTAR REFERENSI	46



DAFTAR TABEL

Tabel 2.1 Daftar Penelitian Terdahulu yang Relevan.....	4
Tabel 2.2 <i>Confusion Matrix</i>	10
Tabel 4.1 Data Latih untuk Manualisasi	21
Tabel 4.2 Hasil Normalisasi Data Latih.....	21
Tabel 4.3 Matriks Bobot untuk $k = 1$	22
Tabel 4.4 Matriks Bias untuk $k = 1$	22
Tabel 4.5 Matriks H_Init untuk $k = 1$	23
Tabel 4.6 Matriks Hasil Keluaran <i>Hidden Layer</i>	23
Tabel 4.7 Matriks $(HTH) - 1$ untuk $k = 1$	24
Tabel 4.8 Matriks H^+ untuk $k = 1$	24
Tabel 4.9 Matriks Beta pada $k = 1$	25
Tabel 4.10 Matriks Bobot pada $k = 2$	25
Tabel 4.11 Matriks Bias pada $k = 2$	25
Tabel 4.12 Matriks Beta pada $k = 2$	25
Tabel 4.13 Matriks Bobot pada $k = 3$	26
Tabel 4.14 Matriks Bias pada $k = 3$	26
Tabel 4.15 Matriks Beta pada $k = 3$	26
Tabel 4.16 Hasil Klasifikasi Data Latih pada $k = 1$	26
Tabel 4.17 Hasil Klasifikasi Data Latih pada $k = 2$	27
Tabel 4.18 Hasil Klasifikasi Data Latih pada $k = 3$	27
Tabel 4.19 Data Uji untuk Manualisasi	28
Tabel 4.20 Hasil Normalisasi Data Uji	29
Tabel 4.21 Matriks H_Init pada $k = 1$	29
Tabel 4.22 Hasil Keluaran <i>Hidden Layer</i>	30
Tabel 4.23 Hasil Klasifikasi Data Uji pada $k = 1$	30
Tabel 4.24 Hasil Klasifikasi Data Uji pada $k = 1$	30
Tabel 4.25 Hasil Klasifikasi Data Uji pada $k = 2$	31
Tabel 4.26 Hasil Klasifikasi Data Uji pada $k = 3$	31
Tabel 4.27 Hasil Klasifikasi Data Uji dari <i>Voting</i> Tiga Prediksi.....	32
Tabel 4.28 Hasil Evaluasi <i>Confusion Matrix</i>	32

Tabel 4.29 Pengujian Jumlah <i>Hidden Neuron</i>	33
Tabel 4.30 Pengujian Nilai <i>k</i>	33
Tabel 4.31 Pengujian Fungsi Aktivasi	34
Tabel 6.1 Hasil Pengujian Banyak <i>Hidden Neuron</i>	40
Tabel 6.2 Pengujian Fungsi Aktivasi	42
Tabel 6.3 Hasil Pengujian Nilai <i>k</i>	43



DAFTAR GAMBAR

Gambar 2.1 Arsitektur ELM.....	7
Gambar 3.1 Diagram Alir Metode <i>Voting based Extreme Learning Machine</i> (Bagian 1)	12
Gambar 3.2 Diagram Alir Metode <i>Voting based Extreme Learning Machine</i> (Bagian 2)	13
Gambar 4.1 Diagram Alir Metode <i>Voting based Extreme Learning Machine</i>	15
Gambar 4.2 Diagram Alir Proses Normalisasi (Bagian 1)	16
Gambar 4.3 Diagram Alir Proses Normalisasi (Bagian 2)	17
Gambar 4.4 Diagram Alir Proses Pelatihan	18
Gambar 4.5 Diagram Alir Proses Pengujian (Bagian 1)	19
Gambar 4.6 Diagram Alir Proses Pengujian (Bagian 2)	20
Gambar 5.1 Antarmuka Hasil Klasifikasi	39
Gambar 5.2 Antarmuka Hasil Evaluasi	39
Gambar 6.1 Grafik Hasil Pengujian Banyak <i>Hidden</i> Neuron	41
Gambar 6.2 Grafik Hasil Pengujian Fungsi Aktivasi	42
Gambar 6.3 Grafik Hasil Pengujian Nilai k	44

BAB 1 PENDAHULUAN

1.1 Latar belakang

Pertumbuhan abnormal dari sel-sel payudara dapat menimbulkan tumor ganas yang disebut kanker payudara (Kementerian Kesehatan, 2016). Menurut data yang dihimpun oleh *World Health Organization* (WHO) (2018), kanker payudara mengakibatkan hingga 2,1 juta perempuan meninggal setiap tahunnya. Diperkirakan pada tahun 2018 kanker payudara mengakibatkan sekitar 627.000 perempuan meninggal dunia. Berdasarkan data Riset Kesehatan Dasar Kementerian Kesehatan RI 2013 (dalam liputan6, 2018), prevalensi kanker tertinggi di Indonesia salah satunya adalah kanker payudara. Penyakit ini menyebabkan perempuan di Indonesia meninggal dunia dengan angka kematian 21,5 per 100.000 penduduk.

Upaya pencegahan dengan *screening* kanker payudara merupakan langkah yang dapat dilakukan untuk mengurangi jumlah kematian akibat kanker payudara (Junaidi, 2014). *Screening* adalah pendeteksian kanker sebelum pasien mengalami gejala (*World Health Organization*, 2018). Perkembangan teknologi medis dan teknologi informasi, di dalam dunia medis dapat digunakan peneliti di bidangnya untuk mengembangkan model deteksi dini (Wang et al., 2018), dari data konsultasi rutin dan analisis darah (Patricio et al., 2018). *Data mining* dapat digunakan untuk mendeteksi penyakit kanker dengan menggunakan teknik klasifikasi, kemudian hasil klasifikasi tersebut digunakan untuk membantu dokter dalam mendiagnosis penyakit (Adinugroho dan Sari, 2018).

Metode *Extreme Learning Machine* (ELM) merupakan salah satu metode yang dapat digunakan untuk melakukan pengklasifikasian dengan komputer. Sebelumnya telah dilakukan penelitian mengenai deteksi dini keberadaan kanker payudara dengan metode *Support Vector Machine* (SVM) yang dilakukan oleh Patricio et al (2018). Penelitian tersebut mendapatkan hasil *sensitivity* 82% hingga 88%, hasil *specificity* 85% hingga 90%. Penelitian terdahulu yang terkait dengan metode ELM untuk masalah pengklasifikasian yang kedua adalah penelitian yang dilakukan oleh Fadilla et al (2018), pada penelitian tersebut memberikan hasil akurasi terbaik yaitu sebesar 96,7% dengan 50 *hidden* neuron. Kemudian yang ketiga adalah penelitian yang dilakukan oleh Cao et al (2012) dengan metode *Voting Based Extreme Learning Machine*, dijelaskan bahwa dalam metode ELM masih terdapat kesalahan dalam melakukan klasifikasi pada sampel yang dekat dengan batas klasifikasi antar kelas, hasil dari penelitian tersebut diperoleh bahwa metode V-ELM mampu mendapatkan tingkat keberhasilan dalam melakukan klasifikasi dibanding metode ELM.

Berdasarkan uraian yang telah dijelaskan di atas serta penelitian yang sebelumnya telah dilakukan, melakukan deteksi dini merupakan hal yang penting untuk mengetahui keberadaan kanker payudara, maka dari itu penulis tertarik untuk melakukan penelitian dengan judul "Klasifikasi Pengidap Kanker Payudara Menggunakan Metode *Voting Based Extreme Learning Machine* (V-ELM)". Aplikasi

yang akan dibangun diharapkan dapat membantu dokter dalam mendiagnosis penyakit kanker payudara.

1.2 Rumusan masalah

Berlandaskan pada latar belakang yang telah diuraikan, maka rumusan masalah yang dibahas pada penelitian ini sebagai berikut.

1. Bagaimana merancang metode *Voting Based Extreme Learning Machine* (V-ELM) dalam mengklasifikasikan pengidap kanker payudara?
2. Berapa nilai parameter yang optimal pada metode *Voting Based Extreme Learning Machine* (V-ELM) untuk klasifikasi pengidap kanker payudara?
3. Bagaimana hasil *sensitivity*, *specificity* dan akurasi yang diperoleh dari penggunaan metode *Voting Based Extreme Learning Machine* (V-ELM) untuk klasifikasi pengidap kanker payudara?

1.3 Tujuan

Tujuan yang ingin diraih dalam penelitian ini, antara lain:

1. Merancang metode *Voting Based Extreme Learning Machine* (V-ELM) untuk mengklasifikasikan pengidap kanker payudara.
2. Mendapatkan nilai parameter yang optimal pada metode *Voting Based Extreme Learning Machine* (V-ELM) untuk klasifikasi pengidap kanker payudara.
3. Mendapatkan hasil *sensitivity*, *specificity* dan akurasi yang diperoleh dari penggunaan metode *Voting Based Extreme Learning Machine* (V-ELM) untuk klasifikasi pengidap kanker payudara.

1.4 Manfaat

Sistem klasifikasi penyakit kanker payudara dibangun dengan harapan dapat memberikan manfaat bagi dokter sebagai alat bantu dalam mendiagnosis penyakit kanker payudara serta mendapatkan nilai parameter terbaik untuk V-ELM dan hasil evaluasi seperti *sensitivity*, *specificity* dan akurasi untuk klasifikasi pengidap kanker payudara.

1.5 Batasan masalah

Untuk menghindari pembahasan yang menyimpang dari rumusan masalah, sehingga penulis membatasi permasalahan, diantaranya sebagai berikut.

1. Data yang digunakan adalah *Breast Cancer Coimbra Dataset* tahun 2018 dari *UCI Machine Learning Repository*.
2. Metode yang digunakan adalah *Voting Based Extreme Learning Machine* (V-ELM).
3. Jumlah kelas yang digunakan dalam melakukan klasifikasi ada 2, yaitu *patient* dan *healthy control*.

1.6 Sistematika Pembahasan

Sistematika pembahasan dalam penelitian ini terdiri dari tujuh bab dengan uraian sebagai berikut.

Bab 1 PENDAHULUAN

Pada bab ini menjelaskan mengenai latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, serta sistematika pembahasan.

Bab 2 LANDASAN KEPUSTAKAAN

Dasar teori serta referensi yang terkait dengan penelitian yang dilakukan akan dijelaskan pada bagian bab ini.

Bab 3 METODOLOGI PENELITIAN

Pada bab ini menjelaskan mengenai metodologi penelitian, diantaranya tipe penelitian yang diterapkan dalam penelitian ini serta strategi dan rancangan penelitian ini.

Bab 4 PERANCANGAN

Pada bab ini menjelaskan tentang formulasi permasalahan dan perancangan strategi penyelesaian masalah.

Bab 5 IMPLEMENTASI

Pada bab ini menjelaskan tentang implementasi yang didasarkan dari hasil dari perancangan pada bab 4.

Bab 6 PENGUJIAN DAN ANALISIS

Pada bab ini menguraikan hasil dari pengujian dan analisis yang dilakukan.

Bab 7 PENUTUP

Pada bab ini menjelaskan tentang kesimpulan yang diperoleh dari hasil penelitian dan saran yang dapat diberikan untuk penelitian selanjutnya.

BAB 2 LANDASAN KEPUSTAKAAN

2.1 Kajian Pustaka

Daftar penelitian terdahulu yang memiliki relevansi terkait dengan penelitian ini ditunjukkan pada Tabel 2.1.

Tabel 2.1 Daftar Penelitian Terdahulu yang Relevan

No	Judul	Objek	Metode	Hasil
1	<i>Using Resistin, glucose, age and BMI to predict the presence of breast cancer</i> (Patricio et al., 2018)	<i>Breast Cancer</i>	<i>Support Vector Machine (SVM)</i>	Hasil <i>sensitivity</i> 82% hingga 88% dan hasil <i>specificity</i> 85% hingga 90%.
2	<i>Classification and Regression of ELM, LVQ and SVM for E-nose Data of Strawberry Juice</i> (Qiu et al., 2015)	<i>E-nose Data of Strawberry Juice</i>	<i>Extreme Learning Machine, Learning Vector Quantization (LVQ), Support Vector Machine (SVM)</i>	Metode ELM memberikan performa terbaik untuk klasifikasi dan regresi dibandingkan metode LVQ dan SVM untuk data <i>E-nose</i> dari jus stroberi.
3	Klasifikasi Penyakit <i>Chronic Kidney Disease</i> (CKD) Dengan menggunakan Metode <i>Extreme Learning Machine</i> (ELM) (Fadilla et al., 2018)	Penyakit <i>Chronic Kidney Disease</i>	<i>Extreme Learning Machine</i> (ELM)	Hasil optimal yang didapatkan yaitu, rata-rata tingkat akurasi 96,7% dan 50 <i>hidden neuron</i> .
4	<i>Voting based Extreme Learning Machine</i> (Cao et al., 2012)		<i>Voting Based Extreme Learning Machine</i>	Dilakukan perbandingan dengan beberapa algoritma klasifikasi, metode <i>V-ELM</i> mampu

Tabel 2.1 Daftar Penelitian Terdahulu yang Relevan (lanjutan)

				memperoleh tingkat keberhasilan yang baik pada beberapa <i>dataset</i> yang dilakukan pengujian. Pada penelitian ini disarankan untuk menggunakan nilai k dalam rentang 5 sampai 35.
--	--	--	--	--

2.2 Payudara

Payudara merupakan salah satu organ pada tubuh yang terdiri dari lemak, jaringan ikat, lobus dan duktus. Fungsi payudara pada wanita berperan untuk menghasilkan air susu ibu (ASI) bagi bayi dan mempunyai peran dalam aktivitas seksual (Alodokter, 2018).

2.2.1 Kanker Payudara

Kanker payudara adalah kondisi ketika terbentuknya sel di payudara secara tidak normal dan tidak terkendali. Bagian pada payudara seperti lobulus, duktus, jaringan lemak atau jaringan ikat dapat terjangkit kanker (Alodokter, 2014).

2.2.2 Faktor Kanker Payudara

Beberapa faktor yang digunakan untuk mendiagnosis kanker payudara antara lain (Patricio et al., 2018):

1. Usia

Jika dirujuk dari Kamus Besar Bahasa Indonesia atau KBBI, usia adalah lama waktu hidup sejak dilahirkan.

2. Indeks Massa Tubuh (IMT)

IMT ialah satuan indeks yang digunakan untuk mengetahui status gizi seseorang dengan melakukan perhitungan antara berat badan dan tinggi badan (Matin dan Veria, 2013).

3. Glukosa

Glukosa darah adalah kandungan gula yang dibentuk dari karbohidrat makanan yang mengalir bersama darah dan tubuh akan menyimpannya ke dalam hati dan otot rangka sebagai glikogen (Joyce, 2006).

4. Resistin

Resistin merupakan hasil sekresi adipositokin oleh jaringan adipose yang mana kadarnya dapat meningkat pada seseorang yang mengalami obesitas (Shuldiner, 2001).

5. Insulin

Menurut Kamus Besar Bahasa Indonesia (KBBI), insulin merupakan pengendali kadar gula dalam darah yang pembentukannya terjadi di pankreas.

6. HOMA

Homeostatic model assessment (HOMA) adalah metode yang digunakan untuk menghitung daya tahan insulin (Patricio et al., 2018).

7. Leptin

Menurut Kamus Besar bahasa Indonesia (KBBI), leptin merupakan hormon yang berfungsi untuk mengontrol perilaku manusia melalui otak untuk makan secukupnya. Leptin dihasilkan dari sel-sel lemak yang kemudian mengalir dengan darah.

8. Adiponektin

Adiponektin merupakan hormon yang berfungsi menghambat nafsu makan, meningkatkan metabolisme tubuh dan dapat memberikan peningkatan terhadap kemampuan otot dalam memanfaatkan karbohidrat (Wolipop, 2012).

9. MCP.1

Monocytes Chemoattractant Protein.1 (MCP.1) adalah *subfamily* kemokin yang berfungsi sebagai *chemoattractant* yang kuat terhadap monosit (Daly et al., 2003).

2.3 Klasifikasi

Klasifikasi adalah proses mengelompokkan data dengan karakteristik yang sama kedalam kelas-kelas yang telah ditentukan terlebih dahulu. Label kelas dalam klasifikasi digunakan untuk menamai kelompok yang memiliki pola yang serupa (Sugianto et al., 2018).

2.4 Normalisasi Data

Normalisasi data masukan dengan menggunakan Persamaan 2.1

$$x'_{ij} = \frac{x_{ij} - \min_j}{\max_j - \min_j} \quad (2.1)$$

Keterangan:

x'_{ij} = nilai hasil normalisasi data ke- i fitur ke- j

x_{ij} = nilai data ke- i fitur ke- j

max_j = nilai terbesar pada fitur ke- j
 min_j = nilai terkecil pada fitur ke- j

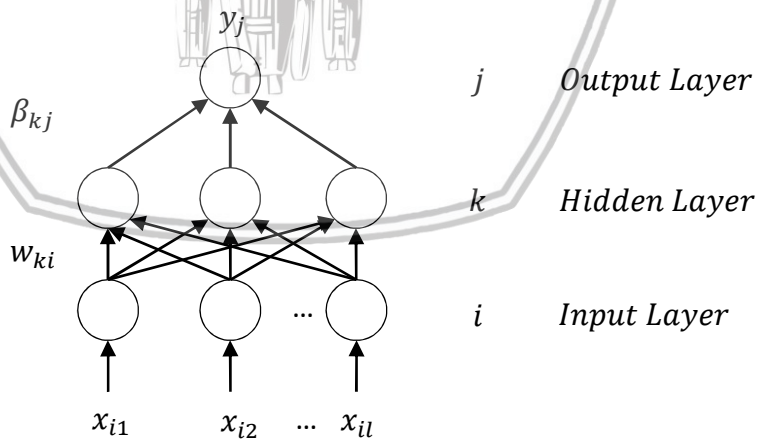
2.5 Jaringan Saraf Tiruan

Jaringan Saraf Tiruan (JST) merupakan suatu pemodelan dengan meniru cara kerja sistem saraf manusia untuk melakukan tugas tertentu (Suyanto, 2014). Menurut Wicaksono (2018), hasil yang didapatkan dari jaringan saraf tiruan merupakan hasil dari proses pelatihan. Pada proses pelatihan masukan yang digunakan untuk jaringan adalah pola-pola data dan target. Pratama (2018) menjelaskan bahwa jaringan saraf tiruan terdiri dari tiga lapisan, antara lain *input layer*, *hidden layer* dan *output layer*. *Input layer* memiliki fungsi untuk menerima fitur sebagai data *input*. *Hidden layer* memiliki fungsi untuk mengekstraksi pola data masukan. *Output layer* memiliki fungsi untuk memberikan hasil keluaran dari *input layer* dan *hidden layer*.

2.6 Extreme Learning Machine (ELM)

Jaringan saraf tiruan yang memiliki satu *hidden layer* yang dapat digunakan untuk melakukan klasifikasi dan regresi adalah metode *Extreme Learning Machine* (ELM). Pada tahap pelatihan metode ELM dilakukan pengacakan untuk inisialisasi nilai bobot dan bias serta menetapkan banyak *hidden* neuron yang akan digunakan (Ding et al., 2015).

2.6.1 Arsitektur ELM



Gambar 2.1 Arsitektur ELM

2.6.2 Fungsi Aktivasi

Pada *Extreme Learning Machine* (ELM) terdapat beberapa fungsi aktivasi yang digunakan untuk meneruskan sinyal dari *input layer* ke *output layer*. Menurut (Huang et al., 2015) beberapa fungsi aktivasi tersebut diantaranya:

1. Fungsi Aktivasi *Sigmoid*

Untuk menghitung fungsi aktivasi *Sigmoid* menggunakan Persamaan 2.2.

$$f(x) = \frac{1}{1+\exp^{-H_Init}} \quad (2.2)$$

2. Fungsi Aktivasi *Fourier*

Untuk menghitung fungsi aktivasi *Fourier* menggunakan Persamaan 2.3.

$$f(x) = \sin(H_Init) \quad (2.3)$$

3. Fungsi Aktivasi *Hyperbolic Tangen*

Untuk menghitung fungsi aktivasi *Hyperbolic Tangen* menggunakan Persamaan 2.4.

$$f(x) = \frac{1-\exp^{-H_Init}}{1+\exp^{-H_Init}} \quad (2.4)$$

4. Fungsi Aktivasi Linear

Untuk menghitung fungsi aktivasi *Linear* menggunakan Persamaan 2.5.

$$f(x) = H_init \quad (2.5)$$

2.6.3 Proses *Training* ELM

1. Inisialisasi nilai *random* W_{jk} untuk input bobot dan nilai bias b_j .

Keterangan:

j = indeks *hidden* neuron

k = indeks masukan

2. Menghitung matriks H_{init} sebagai masukan *hidden layer* dengan Persamaan 2.6.

$$H_{init} = (X * W^T) + b \quad (2.6)$$

Keterangan:

H_{init} = matriks masukan *hidden layer*

X = matriks data masukan

W^T = *Transpose* dari matriks bobot

b = nilai bias

3. Hitung matriks H sebagai matriks keluaran dari *hidden layer* menggunakan Persamaan 2.7.

$$H = \frac{1}{1+\exp(-H_{init})} \quad (2.7)$$

Keterangan:

H = matriks keluaran *hidden layer*

\exp = Eksponensial

H_{init} = matriks masukan *hidden layer*

- Melakukan perhitungan untuk mencari matriks H^+ dengan *Moore-Penrose Generalized Inverse* seperti Persamaan 2.8.

$$H^+ = (H^T H)^{-1} H^T \quad (2.8)$$

Keterangan:

H^+ = matriks *Moore-Penrose*

H^T = *Transpose* dari matriks H

- Melakukan perhitungan untuk mencari matriks Beta sebagai bobot dari *hidden layer* menuju *output layer* dengan Persamaan 2.9.

$$\hat{\beta} = H^+ Y \quad (2.9)$$

Keterangan:

$\hat{\beta}$ = matriks *output weight*

H^+ = matriks *Moore-Penrose*

Y = matriks target

2.6.4 Proses *Testing* ELM

- Melakukan perhitungan pada Persamaan 2.6 sampai dengan Persamaan 2.7 dengan menggunakan matriks bobot w , bias b dan $\hat{\beta}$ dari hasil proses *training*.
- Melakukan perhitungan untuk mencari hasil *output layer* atau hasil klasifikasi, dengan Persamaan 2.10.

$$\hat{Y} = H \hat{\beta} \quad (2.10)$$

Keterangan:

$\hat{\beta}$ = matriks *output weight*

H = matriks fungsi aktivasi

\hat{Y} = matriks *output* hasil klasifikasi

2.7 *Voting Based Extreme Learning Machine* (V-ELM)

Salah satu masalah yang dimiliki oleh metode ELM ialah kemungkinan terjadi kesalahan klasifikasi pada sampel yang berada dekat dengan batas klasifikasi antar kelas. Permasalahan tersebut dapat diatasi dengan menerapkan metode *voting based* pada ELM (V-ELM) (Ginting et al., 2018).

Pada metode V-ELM akan melakukan perulangan metode ELM sebanyak k kali. Untuk setiap ELM menggunakan learning parameter yang berbeda, seperti bobot dan bias. Sedangkan untuk data masukan menggunakan data latih yang sama untuk tiap ELM. Kemudian untuk menentukan kelas dari data ditentukan dengan *voting* terbanyak dari setiap ELM.

2.7.1 Algoritme V-ELM

Tahap *training* V-ELM:

1. Inisialisasi nilai k sebagai banyaknya perulangan ELM tahap *training*, bernilai ganjil.
2. Selama banyaknya perulangan yang dilakukan masih dalam rentang nilai k , ulangi langkah 3-5.
3. Inisialisasi bobot w dan bias b secara acak.
4. Melakukan perhitungan dengan Persamaan 2.6 sampai dengan Persamaan 2.9.
5. Menambah nilai k dengan 1.

Tahap *testing* V-ELM:

1. Inisialisasi nilai k sebagai banyaknya perulangan ELM tahap *testing*, bernilai ganjil.
2. Selama banyaknya perulangan yang dilakukan masih dalam rentang nilai k , ulangi langkah 3-5.
3. Dengan parameter hasil *training* (w_j^k, b_j^k, β_j^k), lakukan proses *testing* ELM.
4. Simpan hasil keluaran klasifikasi ke dalam matriks $s_{k,x}$, yang mana x = data *input*.
5. Menambah nilai k dengan 1.
6. Hasil klasifikasi adalah kelas dengan *voting* tertinggi untuk tiap data berdasarkan matriks $s_{k,x}$.

2.8 Nilai Evaluasi

Evaluasi dilakukan untuk mengetahui kemampuan dari metode V-ELM dalam melakukan klasifikasi penderita kanker payudara. Kinerja dari sistem klasifikasi yang telah dibuat, selanjutnya dilakukan pengukuran menggunakan *confusion matrix*. Pada *confusion matrix* terdapat beberapa istilah seperti *sensitivity*, *specificity*, akurasi (Adinugroho dan Sari, 2018). Tabel 2.2 mendeskripsikan *confusion matrix*.

Tabel 2.2 Confusion Matrix

Nilai prediksi	Nilai yang sebenarnya	
	TRUE	FALSE
TRUE	TP (<i>True Positive</i>)	FP (<i>False Positive</i>)
FALSE	FN (<i>False Negative</i>)	TN (<i>True Negative</i>)

1. *True Positive* (TP) adalah hasil prediksi bernilai benar sesuai dengan data yang bernilai benar.
2. *False Positive* (FP) adalah hasil prediksi bernilai benar, lalu data yang sebenarnya bernilai salah.

3. *False Negative* (FN) adalah hasil prediksi bernilai salah, sementara data yang sebenarnya bernilai benar.
4. *True Negative* (TN) adalah hasil prediksi bernilai salah dan data yang sebenarnya bernilai salah.

Perhitungan yang dilakukan dalam metode *confusion matrix* terdapat 4, antara lain:

1. *Sensitivity* adalah hasil perbandingan antara jumlah data yang diklasifikasikan benar oleh sistem dengan jumlah data yang benar. Persamaan 2.11 digunakan untuk menghitung *Sensitivity*.

$$Sensitivity = \frac{TP}{TP+FN} \quad (2.11)$$

2. *Specificity* adalah proses perhitungan hasil prediksi salah yang benar dibagi dengan semua data yang bernilai salah Persamaan 2.12 digunakan untuk menghitung *Specificity*.

$$Specificity = \frac{TN}{TN+FP} \quad (2.12)$$

3. *Akurasi* adalah proses perhitungan seluruh nilai prediksi yang benar dibagi dengan banyak data yang diklasifikasikan. Persamaan 2.13 digunakan untuk menghitung Akurasi.

$$Akurasi = \frac{TP+TN}{TP+TN+FN+FP} \quad (2.13)$$

BAB 3 METODOLOGI

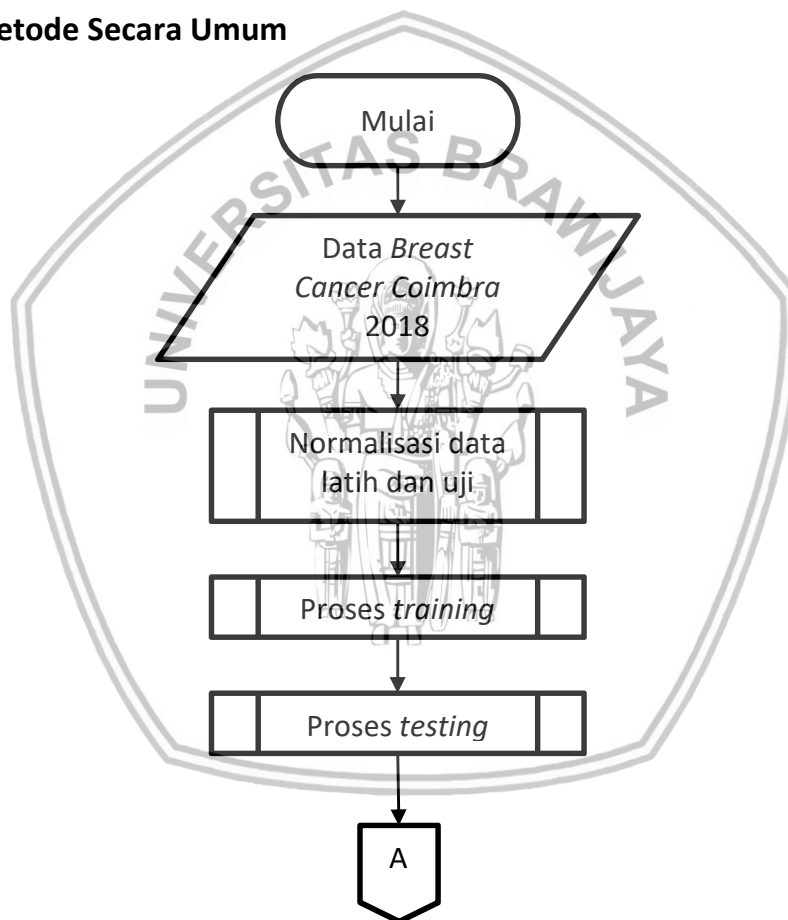
3.1 Tipe penelitian

Penulis melakukan penelitian dengan tipe non-implementatif analitik, yang mana tujuannya ialah menganalisis data kanker payudara yang menjadi objek dengan menggunakan metode V-ELM.

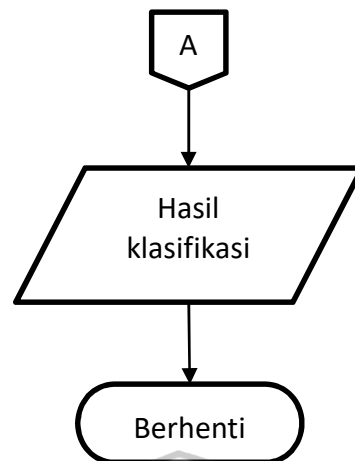
3.2 Strategi dan Rancangan Penelitian

Penjelasan mengenai strategi dan rancangan penelitian yang dilakukan oleh penulis dijabarkan sebagai berikut.

3.2.1 Metode Secara Umum



Gambar 3.1 Diagram Alir Metode *Voting based Extreme Learning Machine* (Bagian 1)



Gambar 3.2 Diagram Alir Metode *Voting based Extreme Learning Machine* (Bagian 2)

Secara umum, metode yang dilakukan pada penelitian ini ditunjukkan pada Gambar 3.1 dan Gambar 3.2. Masukan yang dibutuhkan untuk penelitian ini adalah berupa *dataset Breast Cancer Coimbra* yang telah dibagi menjadi data latih dan data uji dengan perbandingan 80:20. Sebelum melanjutkan ke dalam metode *Voting based Extreme Learning Machine* (V-ELM), data tersebut dilakukan normalisasi terlebih dahulu. Setelah itu, akan dilakukan proses pelatihan V-ELM terhadap data latih sehingga menghasilkan bobot, bias dan beta. Kemudian pengujian dilakukan dengan masukan berupa data uji serta bobot, bias dan beta dari hasil pelatihan. Keluaran yang dihasilkan dari penelitian ini adalah berupa label hasil klasifikasi.

3.2.2 Lokasi Penelitian

Penulis melakukan penelitian di Laboratorium Riset Fakultas Ilmu Komputer Universitas Brawijaya.

3.2.3 Metode Pengumpulan Data

Dalam melakukan penelitian, penulis menggunakan data sekunder, yaitu data *Breast Cancer Coimbra Dataset* tahun 2018 yang didapatkan dari *UCI Machine Learning Repository* dengan total data sebanyak 116 data. Atribut dari data yang akan digunakan sebagai fitur untuk masukan sistem dalam melakukan klasifikasi kanker payudara adalah usia, Indeks Massa Tubuh (IMT), glukosa, resistin, HOMA, insulin, leptin, *adiponectin* dan MCP.1. Label yang digunakan untuk melakukan klasifikasi ada 2, yaitu 1 sebagai *healthy control* dan 2 sebagai pasien.

3.2.4 Metode Analisis Data

Metode analisis data yang digunakan ialah *confusion matrix*. Dari *confusion matrix* kemudian dilakukan analisis dengan 4 evaluasi, yaitu *specificity*, akurasi, *sensitivity*. Keempat jenis evaluasi tersebut digunakan untuk mengetahui seberapa baik hasil dari klasifikasi menggunakan V-ELM. Perhitungan untuk

sensitivity menggunakan Persamaan 2.11, untuk *specificity* menggunakan Persamaan 2.12 dan akurasi menggunakan Persamaan 2.13.

3.2.5 Peralatan Pendukung

Peralatan *hardware* dan *software* yang digunakan oleh penulis untuk melakukan penelitian ini diantaranya:

1. Laptop Asus A455L
2. RAM 4 GB
3. Bahasa pemrograman Java
4. IDE Neatbeans 8.1
5. Library JAMA 1.0.3
6. *Operating System* Windows 8 64-bit
7. Data yang digunakan dalam proses training dan testing berformat .csv



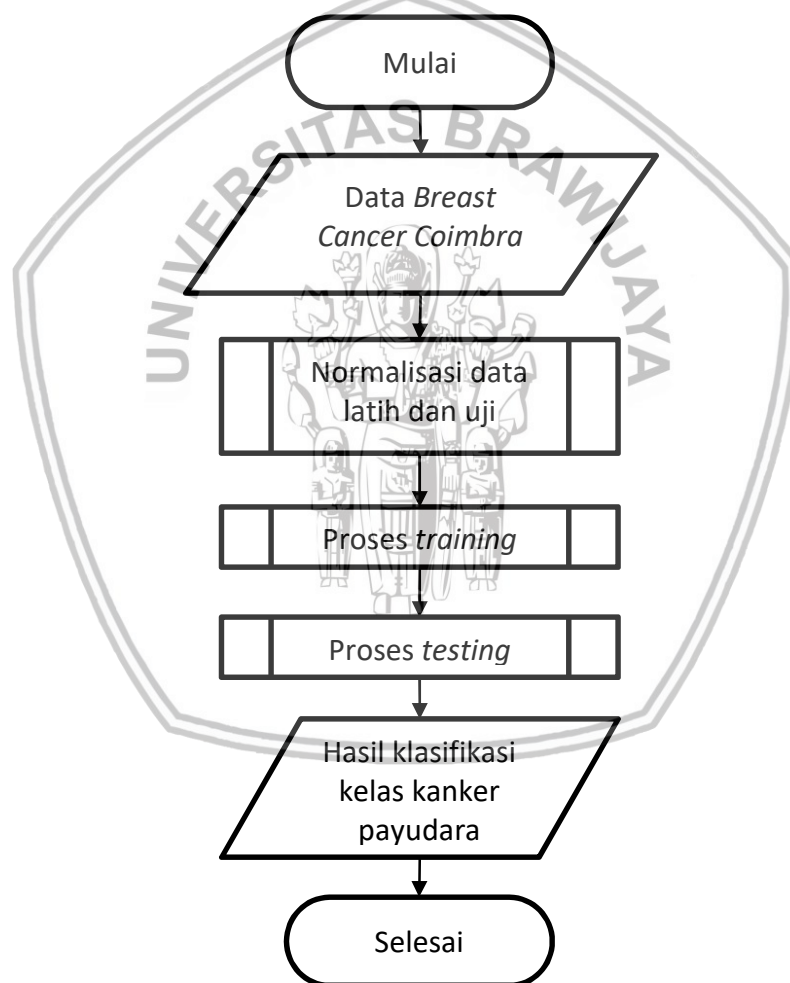
BAB 4 PERANCANGAN

4.1 Formulasi Permasalahan

Pada penelitian ini penulis menggunakan metode V-ELM untuk menyelesaikan masalah klasifikasi pengidap kanker payudara. Data kanker payudara yang akan digunakan untuk melakukan klasifikasi adalah *Breast Cancer Coimbra Data Set* tahun 2018 yang didapatkan dari *UCI Machine Learning Repository*.

4.2 Penyelesaian Masalah

Gambaran secara umum metode V-ELM seperti pada Gambar 4.1.

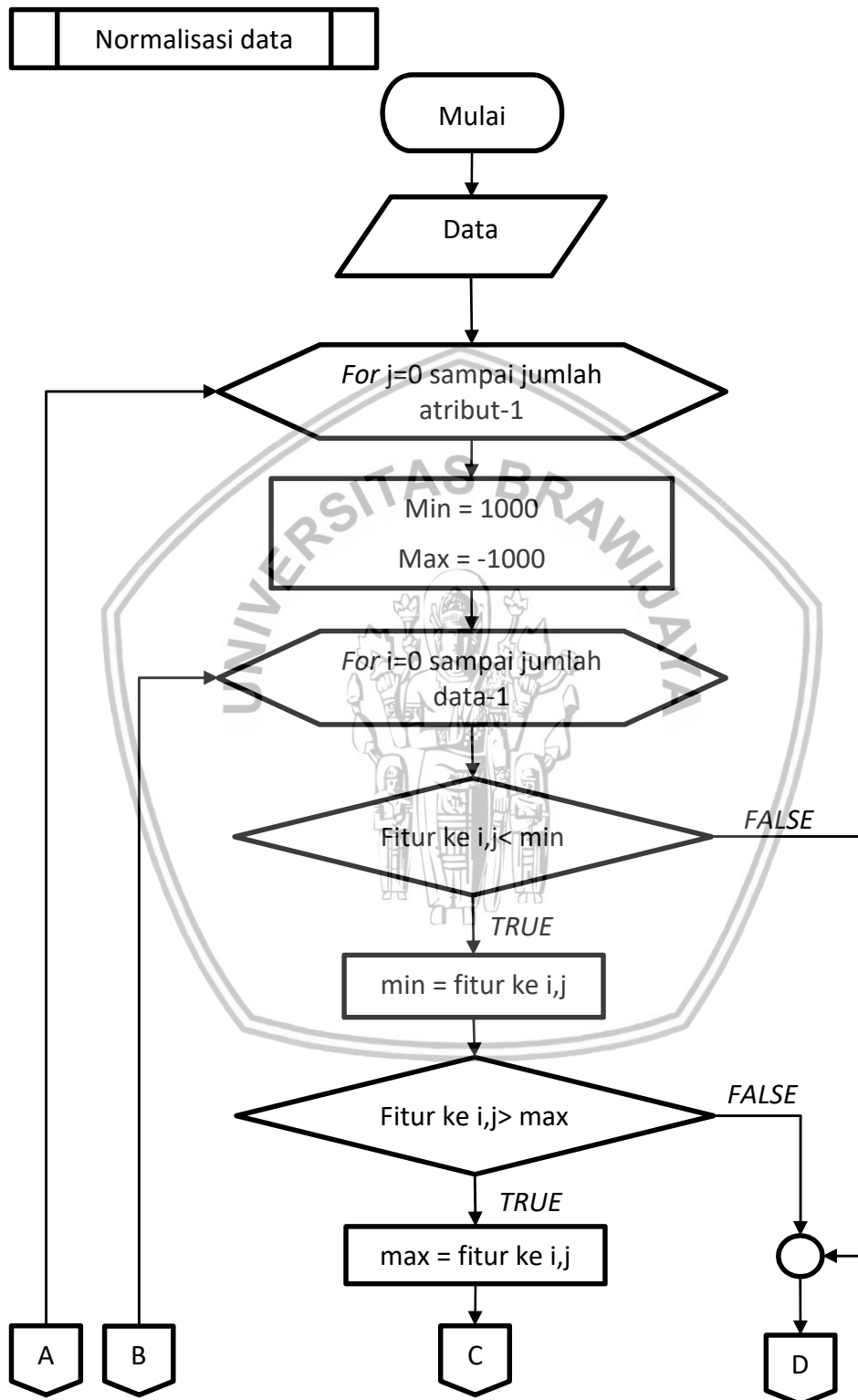


Gambar 4.1 Diagram Alir Metode *Voting based Extreme Learning Machine*

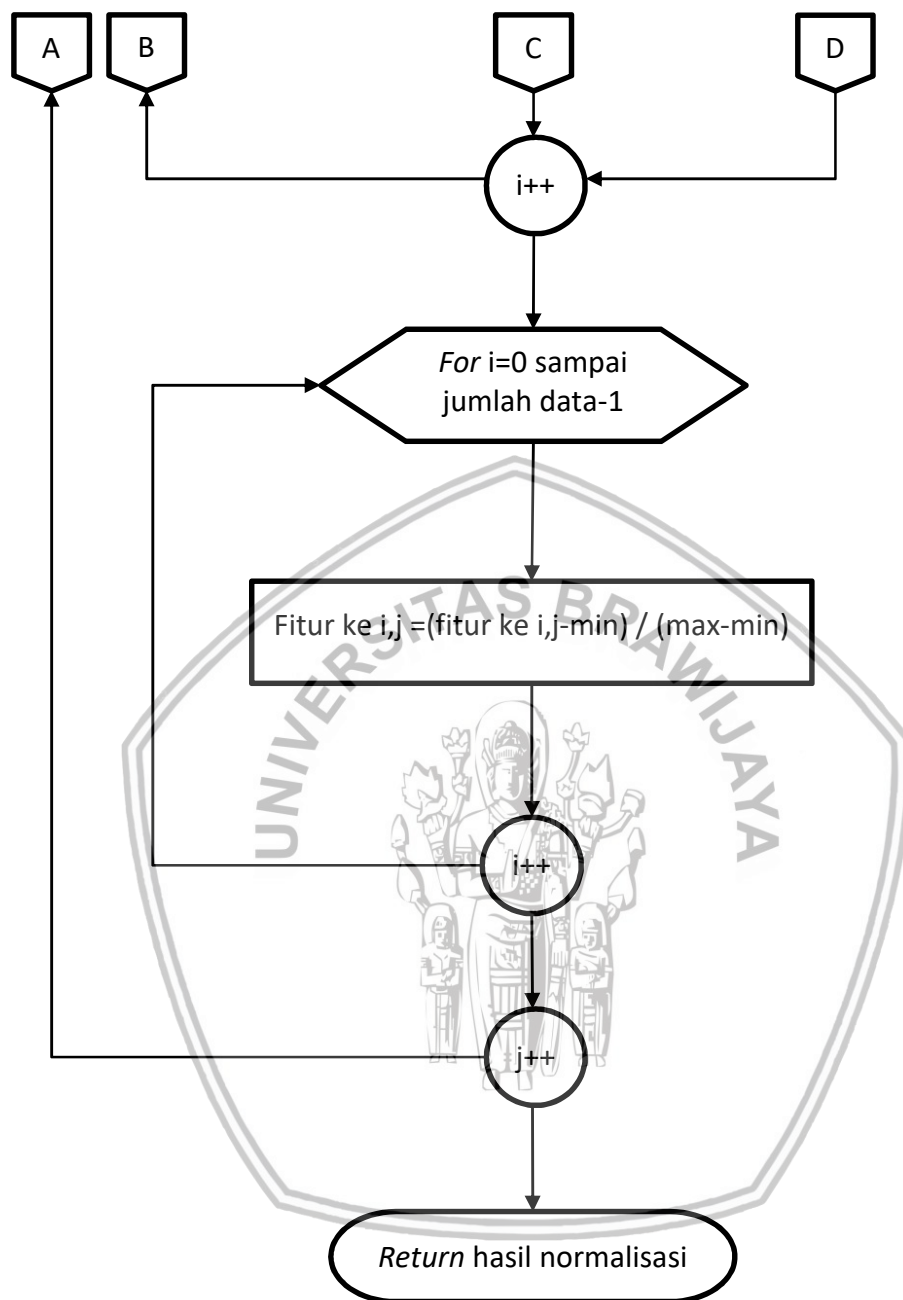
4.2.1 Normalisasi Data

Normalisasi dilakukan untuk mereduksi rentang nilai dari data yang digunakan dalam penelitian ini. Pada penelitian ini digunakan metode *Min-Max*

Normalization digunakan untuk melakukan proses normalisasi. Diagram alir dari proses normalisasi akan ditampilkan pada Gambar 4.2.



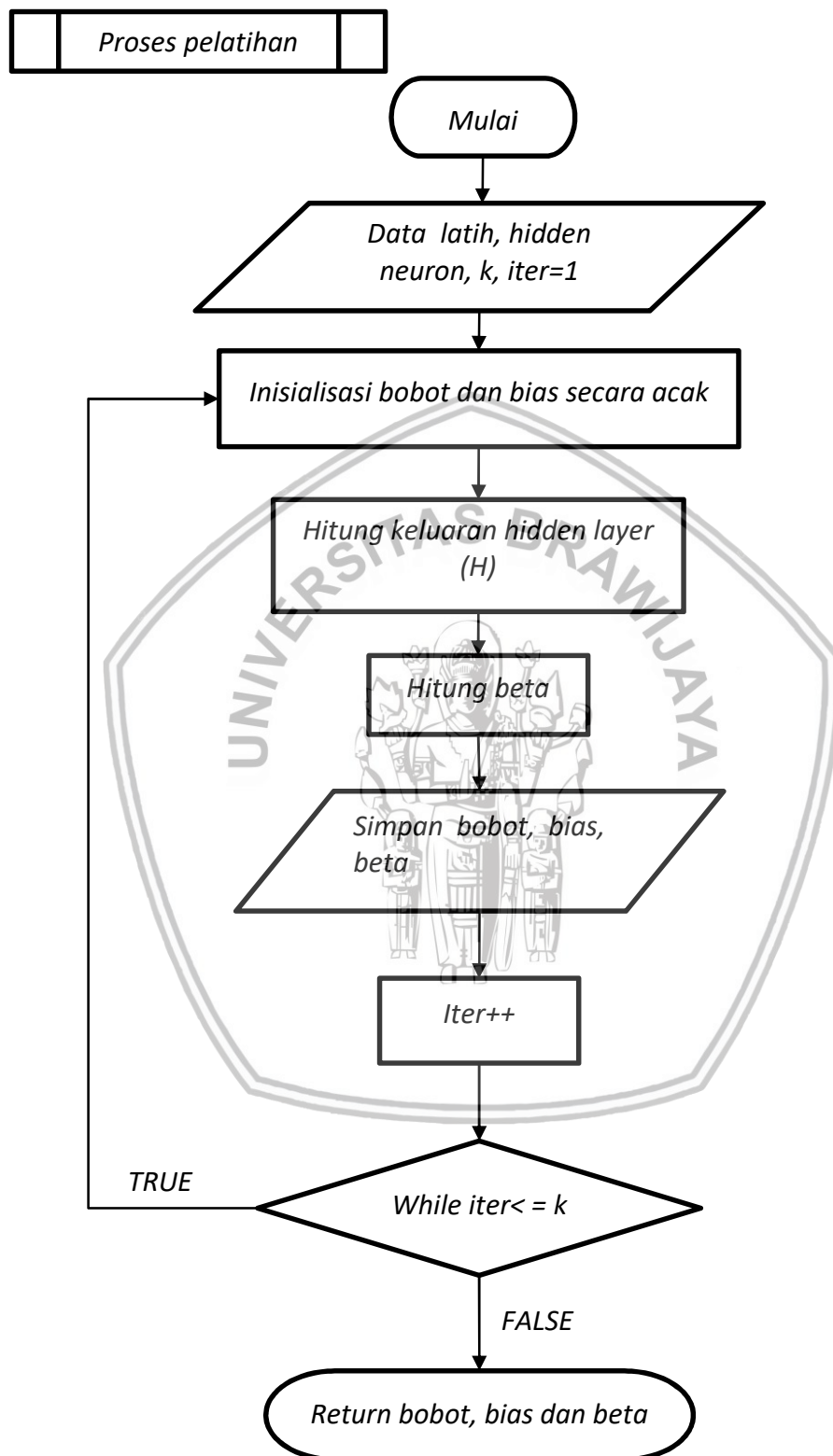
Gambar 4.2 Diagram Alir Proses Normalisasi (Bagian 1)



Gambar 4.3 Diagram Alir Proses Normalisasi (Bagian 2)

4.2.2 Proses Pelatihan V-ELM

Data latih yang telah dinormalisasi seperti pada Gambar 4.3, selanjutnya digunakan sebagai parameter untuk melakukan proses pelatihan. Hasil dari proses pelatihan adalah untuk mendapatkan bobot. Diagram alir dari proses pelatihan metode V-ELM akan ditampilkan pada Gambar 4.4.



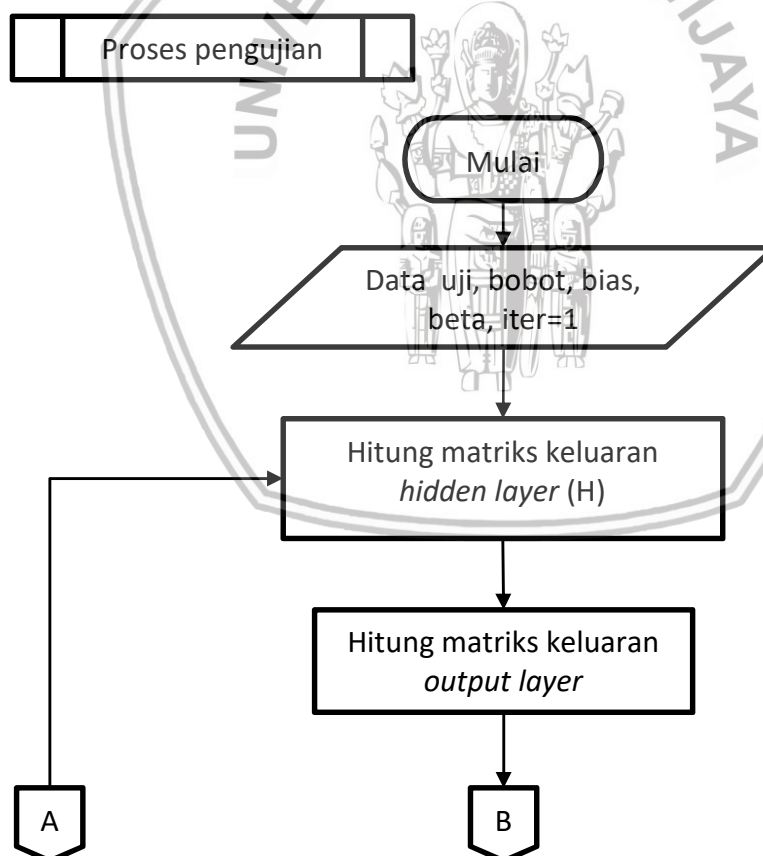
Gambar 4.4 Diagram Alir Proses Pelatihan

Pada proses pelatihan, parameter yang digunakan sebagai masukan diantaranya adalah data latih, banyak *hidden* neuron dan k atau banyak iterasi VELM untuk dilakukan *voting*. Berikut tahapan proses pelatihan pada V-ELM.

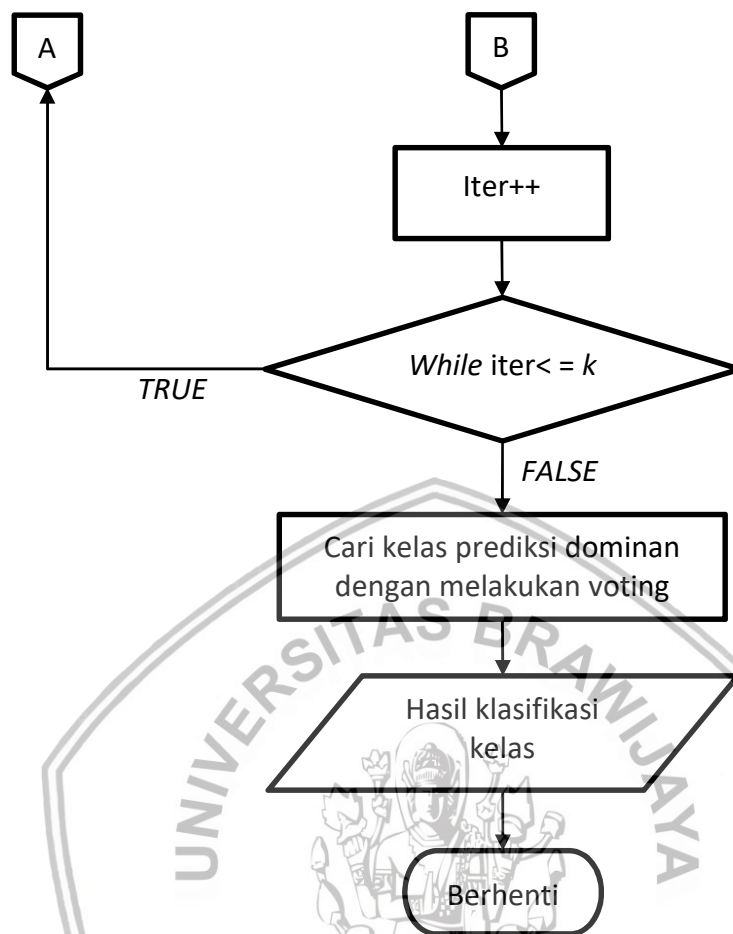
1. Melakukan langkah 2 sampai 5 dengan iterasi sejumlah k .
2. Menginisialisasi bobot dan bias secara acak.
3. Menghitung keluaran dari *hidden layer* menggunakan Persamaan 2.7.
4. Menghitung nilai dari H^+ dengan Persamaan 2.8, kemudian menghitung nilai beta dengan Persamaan 2.9.
5. Menyimpan hasil bobot, bias dan beta untuk setiap perulangan k .

4.2.3 Proses Pengujian V-ELM

Data uji yang telah dinormalisasi seperti pada Gambar 4.3, selanjutnya digunakan untuk melakukan proses pengujian. Parameter yang digunakan pada proses pengujian adalah matriks bobot masukan, matriks bobot bias dan nilai β yang telah didapatkan dari proses sebelumnya, yaitu proses pelatihan. Diagram alir dari proses pengujian metode V-ELM akan ditampilkan di Gambar 4.5 dan Gambar 4.6.



Gambar 4.5 Diagram Alir Proses Pengujian (Bagian 1)



Gambar 4.6 Diagram Alir Proses Pengujian (Bagian 2)

Pada proses pengujian, parameter yang dibutuhkan sebagai masukan diantaranya data uji dan bobot, bias, beta dari proses pelatihan. Berikut penjelasan mengenai diagram alir proses pengujian.

1. Menghitung keluaran dari *hidden layer* dengan menggunakan data uji serta bobot dan bias dari proses pelatihan. Persamaan untuk mendapatkan matriks H sebagai keluaran *hidden layer* adalah Persamaan 2.7.
2. Melakukan perhitungan untuk matriks keluaran dari *output layer* dengan menggunakan Persamaan 2.10 dan nilai beta yang dari hasil pelatihan.
3. Menyimpan matriks Y yang menyimpan hasil keluaran dari *output layer*.
4. Menentukan kelas untuk setiap data dengan mencari nilai maksimum pada setiap *output* neuron.
5. Mengulangi langkah 1 sampai 4 sebanyak iterasi sejumlah nilai k .
6. Melakukan *voting* terhadap hasil kelas setiap data dari k iterasi untuk menentukan kelas prediksi dari setiap data.

4.3 Perhitungan Manualisasi

Perhitungan manual merupakan langkah matematis cara kerja dari metode V-ELM.

4.3.1 Perhitungan Manual Proses Pelatihan algoritme V-ELM

1. Menyiapkan data latih.

Perhitungan manual tahap pelatihan V-ELM pada penelitian ini digunakan sampel sebanyak 9 data dengan 4 fitur. Data untuk perhitungan manualisasi proses pelatihan sebagai berikut.

Tabel 4.1 Data Latih untuk Manualisasi

USIA	IMT	GLUKOSA	RESISTIN	KELAS
35	30,277	84	16,437	1
54	30,483	90	10,193	1
45	37,036	83	8,704	1
50	38,579	106	11,784	1
66	31,447	90	23,382	1
35	35,251	90	22,037	1
55	31,975	92	7,165	2
43	31,25	103	38,653	2
86	26,667	201	24,37	2

Dengan kelas 1 adalah *Healthy Control* dan kelas 2 adalah Pasien.

2. Menormalisasi nilai data dengan memakai Persamaan 2.1.

Nilai hasil normalisasi data ke-1 pada fitur ke-1 (usia):

$$x'_{11} = \frac{35 - 35}{86 - 35} = 0$$

Hasil normalisasi dari data ke-1 pada fitur ke-2 (IMT):

$$x'_{12} = \frac{30,277 - 26,667}{38,579 - 26,667} = 0,303$$

Berikut hasil normalisasi data latih.

Tabel 4.2 Hasil Normalisasi Data Latih

USIA	IMT	GLUKOSA	RESISTIN
0	0,303	0,008	0,294
0,373	0,32	0,059	0,096
0,196	0,87	0	0,049
0,294	1	0,195	0,147
0,608	0,401	0,059	0,515
0	0,721	0,059	0,472

Tabel 4.2 Hasil Normalisasi Data Latih (lanjutan)

USIA	IMT	GLUKOSA	RESISTIN
0,392	0,446	0,076	0
0,157	0,385	0,169	1
1	0	1	0,546

3. Inisialisasi nilai k , jumlah *hidden* neuron.

Pada manualisasi ini nilai k yang digunakan adalah 3 dan jumlah *hidden* neuron yang digunakan adalah 5.

4. Set nilai $k = 1$.

5. Inisialisasi bobot awal secara acak dalam rentang nilai $[-1, 1]$ dengan dimensi baris sebanyak jumlah *hidden* neuron dan dimensi kolom sebanyak jumlah fitur data.

Berikut adalah matriks bobot untuk $k = 1$.

Tabel 4.3 Matriks Bobot untuk $k = 1$

	X1	X2	X3	X4
H1	0,08	0,618	0,044	-0,418
H2	-0,189	-0,336	-0,829	0,937
H3	0,149	0,627	-0,269	-0,665
H4	-0,997	-0,228	-0,37	-0,445
H5	-0,872	-0,689	-0,73	0,918

Dengan H1 adalah bobot menuju ke *hidden* neuron ke-1, begitu pula dengan H2 sampai H5. Kemudian X1 sampai X4 adalah fitur ke-1 sampai ke-5. X1 = usia, X2 = Indeks Massa Tubuh, X3 = glukosa dan X4 = Resistin.

6. Inisialisasi bias awal secara acak dalam rentang nilai $[0, 1]$ dengan dimensi sejumlah *hidden* neuron.

Berikut adalah matriks bias untuk $k = 1$.

Tabel 4.4 Matriks Bias untuk $k = 1$

0,358	0,804	0,126	0,08	0,209
-------	-------	-------	------	-------

7. Menghitung nilai matriks H_Init dengan Persamaan 2.6.

Nilai H_Init untuk data ke-1 pada *hidden* neuron ke-1:

$$H_Init_{11} = (0 \times 0,08) + (0,303 \times 0,618) + (0,008 \times 0,044) + (0,294 \times (-0,418)) + 0,358 = 0,423$$

Berikut matriks H_Init yang didapatkan untuk $k=1$.

Tabel 4.5 Matriks H_Init untuk $k = 1$

0,423	0,971	0,118	-0,123	0,264
0,549	0,667	0,303	-0,429	-0,292
0,891	0,52	0,669	-0,336	-0,517
0,947	0,388	0,647	-0,579	-0,745
0,442	0,988	0,111	-0,869	-0,169
0,609	0,955	0,248	-0,316	0,102
0,669	0,517	0,444	-0,441	-0,496
0,198	1,442	-0,319	-0,672	0,6
0,255	0,298	-0,356	-1,53	-0,892

8. Menghitung matriks H sebagai keluaran dari *hidden layer* dengan memakai Persamaan 2.7.

Berikut nilai keluaran dari data ke-1 pada *hidden* neuron ke-1.

$$H_{11} = \frac{1}{1 + \exp(-0,423)} = 0,604$$

Berikut nilai keluaran dari data ke-1 pada *hidden* neuron ke-2.

$$H_{12} = \frac{1}{1 + \exp(-0,971)} = 0,725$$

Berikut matriks H yang merupakan keluaran dari *hidden layer*.

Tabel 4.6 Matriks Hasil Keluaran *Hidden Layer*

0,604	0,725	0,53	0,469	0,566
0,634	0,661	0,575	0,394	0,427
0,709	0,627	0,661	0,417	0,373
0,721	0,596	0,656	0,359	0,322
0,609	0,729	0,528	0,296	0,458
0,648	0,722	0,562	0,422	0,525
0,661	0,626	0,609	0,392	0,378
0,549	0,809	0,421	0,338	0,646
0,563	0,574	0,412	0,178	0,291

9. Menghitung nilai H^+ dengan menggunakan Persamaan 2.8.

Nilai hasil $(H^T H)$ baris ke-1 kolom ke-1 adalah sebagai berikut.

$$\begin{aligned}(H^T H)_{11} &= (0,604 \times 0,604) + (0,634 \times 0,634) + (0,709 \times 0,709) \\ &\quad + (0,721 \times 0,721) + (0,609 \times 0,609) + (0,648 \times 0,648) \\ &\quad + (0,661 \times 0,661) + (0,549 \times 0,549) + (0,563 \times 0,563) \\ &= 3,365\end{aligned}$$

Berikut hasil invers matriks $(H^T H)^{-1}$ untuk $k = 1$.

Tabel 4.7 Matriks $(H^T H)^{-1}$ untuk $k = 1$

371,486	-291,534	-216,902	-43,9119	218,787
-291,534	566,386	-118,462	341,204	-577,269
-216,902	-118,462	385,707	-257,446	221,39
-43,9119	341,204	-257,446	332,656	-408,874
218,787	-577,269	221,39	-408,874	625,38

Nilai hasil perkalian $(H^T H)^{-1}$ dengan H^T baris ke-1 kolom ke-1 untuk $k = 1$ adalah sebagai berikut.

$$\begin{aligned}(H^T H)^{-1} H^T_{11} &= (371,486 \times 0,604) + ((-291,534) \times 0,725) \\ &\quad + ((-216,902) \times 0,53) + ((-43,9119) \times 0,469) \\ &\quad + (218,787 \times 0,566) = 1,233\end{aligned}$$

Berikut matriks H^+ untuk $k = 1$.

Tabel 4.8 Matriks H^+ untuk $k = 1$

1,233	-5,75	0,56	6,222	-13,508	4,642	-3,544	3,425	8,341
5,591	9,139	-3,17	-13,583	9,224	-5,812	5,032	-9,362	5,022
-8,299	-0,776	2,285	5,103	10,429	-1,537	0,324	3,323	-12,755
9,477	5,948	-1,386	-9,441	-2,857	-1,267	3,4	-8,104	5,421
-7,47	-9,351	2,573	13,426	-5,104	5,416	-5,536	12,114	-7,905
1,233	-5,75	0,56	6,222	-13,508	4,642	-3,544	3,425	8,341
5,591	9,139	-3,17	-13,583	9,224	-5,812	5,032	-9,362	5,022
-8,299	-0,776	2,285	5,103	10,429	-1,537	0,324	3,323	-12,755

10. Menghitung nilai $\hat{\beta}$ dengan menggunakan Persamaan 2.9.

Nilai hasil $\hat{\beta}$ baris ke-1 kolom ke-1 untuk $k = 1$ adalah sebagai berikut.

$$\begin{aligned}(\beta)_{11} &= (1,233 \times 1) + ((-5,75) \times 1) + (0,56 \times 1) + (6,222 \times 1) + \\ &\quad + ((-13,508) \times 1) + (4,642 \times 1) + ((-3,544) \times 0) + (3,425 \times 0) + (8,341 \times 0) \\ &= -6,601\end{aligned}$$

Berikut matriks $\hat{\beta}$ untuk $k = 1$.

Tabel 4.9 Matriks Beta pada $k = 1$

-6,601	8,222
1,389	0,692
7,204	-9,108
0,474	0,717
-0,51	-1,327

11. Set $k = k + 1$.
12. Ulangi langkah ke-5 sampai dengan ke-11 sampai nilai k terpenuhi.
13. Setelah dilakukan pelatihan sebanyak $k = 3$ didapatkan hasil pelatihan sebagai berikut.

Matriks bobot untuk $k = 2$.

Tabel 4.10 Matriks Bobot pada $k = 2$

	X1	X2	X3	X4
H1	-0,669	-0,92	0,562	0,449
H2	-0,984	0,401	-0,361	0,097
H3	-0,128	-0,354	-0,04	-0,942
H4	0,075	0,069	0,088	0,32
H5	0,817	0,376	0,837	-0,667

Matriks bias untuk $k = 2$.

Tabel 4.11 Matriks Bias pada $k = 2$

0,18	0,176	0,758	0,291	0,66
------	-------	-------	-------	------

Matriks $\hat{\beta}$ untuk $k = 2$.

Tabel 4.12 Matriks Beta pada $k = 2$

-3,013	2,93
0,381	-0,34
2,049	-1,604
3,393	-2,051
-1,938	1,861

Matriks bobot untuk $k = 3$.

Tabel 4.13 Matriks Bobot pada $k = 3$

	X1	X2	X3	X4
H1	0,829	0,151	-0,356	0,725
H2	0,598	-0,513	-0,326	-0,236
H3	0,278	0,248	0,258	-0,029
H4	-0,831	0,993	0,656	0,71
H5	-0,297	0,471	-0,293	0,379

Matriks bias untuk $k = 3$.

Tabel 4.14 Matriks Bias pada $k = 3$

0,541	0,691	0,231	0,384	0,839
-------	-------	-------	-------	-------

Matriks $\hat{\beta}$ untuk $k = 3$.

Tabel 4.15 Matriks Beta pada $k = 3$

-2,355	2,106
-4,1	5,086
5,162	-5,106
-5,534	6,085
7,777	-7,582

14. Menghitung nilai \hat{Y} dengan menggunakan Persamaan 2.10.

Nilai matriks \hat{Y} untuk $k = 1$.

Tabel 4.16 Hasil Klasifikasi Data Latih pada $k = 1$

Data Ke	Y1	Y2	Max	Kelas	Label Prediksi	Label Aktual
1	0,769	0,232	0,769	1	Healthy control	Healthy control
2	0,847	0,145	0,847	1	Healthy control	Healthy control
3	0,961	0,045	0,961	1	Healthy control	Healthy control
4	0,807	0,188	0,807	1	Healthy control	Healthy control
5	0,701	0,308	0,701	1	Healthy control	Healthy control

Tabel 4.16 Hasil Klasifikasi Data Latih pada $k = 1$ (lanjutan)

Data Ke	Y1	Y2	Max	Kelas	Label Prediksi	Label Aktual
6	0,707	0,313	0,707	1	Healthy control	Healthy control
7	0,887	0,1	0,887	1	Healthy control	Pasien
8	0,361	0,628	0,628	2	Pasien	Pasien
9	-0,017	1,018	1,018	2	Pasien	Pasien

Nilai matriks \hat{Y} untuk $k = 2$.

Tabel 4.17 Hasil Klasifikasi Data Latih pada $k = 2$

Data Ke	Y1	Y2	Max	Kelas	Label Prediksi	Label Aktual
1	0,691	0,31	0,691	1	Healthy control	Healthy control
2	0,738	0,262	0,738	1	Healthy control	Healthy control
3	1,015	-0,015	1,015	1	Healthy control	Healthy control
4	0,922	0,078	0,922	1	Healthy control	Healthy control
5	0,676	0,325	0,676	1	Healthy control	Healthy control
6	0,79	0,211	0,79	1	Healthy control	Healthy control
7	0,812	0,187	0,812	1	Healthy control	Pasien
8	0,43	0,57	0,57	2	Pasien	Pasien
9	-0,072	1,072	1,072	2	Pasien	Pasien

Nilai matriks \hat{Y} untuk $k = 3$.

Tabel 4.18 Hasil Klasifikasi Data Latih pada $k = 3$

Data Ke	Y1	Y2	Max	Kelas	Label Prediksi	Label Aktual
1	0,712	0,291	0,712	1	Healthy control	Healthy control

Tabel 4.18 Hasil Klasifikasi Data Latih pada $k = 3$ (lanjutan)

2	0,762	0,234	0,762	1	Healthy control	Healthy control
3	0,964	0,039	0,964	1	Healthy control	Healthy control
4	0,919	0,082	0,919	1	Healthy control	Healthy control
5	0,605	0,398	0,605	1	Healthy control	Healthy control
6	0,812	0,184	0,812	1	Healthy control	Healthy control
7	0,82	0,178	0,82	1	Healthy control	Pasien
8	0,46	0,539	0,539	2	Pasien	Pasien
9	-0,052	1,052	1,052	2	Pasien	Pasien

4.3.2 Perhitungan Proses Pengujian algoritme V-ELM

1. Menyiapkan data uji.

Perhitungan manual tahap pengujian V-ELM pada penelitian ini digunakan sampel data sebanyak 6 dengan 4 fitur. Data untuk perhitungan manualisasi proses pengujian sebagai berikut.

Tabel 4.19 Data Uji untuk Manualisasi

USIA	IMT	GLUKOSA	RESISTIN	KELAS
64	34,53	95	6,702	1
75	23	83	7,091	1
34	21,47	78	6,92	1
71	25,51	112	42,745	2
42	29,297	98	53,672	2
65	29,667	85	19,463	2

Dengan kelas 1 adalah *Healthy Control* dan kelas 2 adalah Pasien.

2. Menormalisasi nilai data dengan memakai Persamaan 2.1.

Nilai hasil normalisasi data ke-1 pada fitur ke-1 (usia):

$$x'_{11} = \frac{64 - 34}{75 - 34} = 0,731$$

Nilai hasil normalisasi data ke-1 pada fitur ke-2 (IMT):

$$x'_{12} = \frac{34,53 - 21,47}{34,53 - 21,47} = 1$$

Berikut hasil data uji yang telah dinormalisasi.

Tabel 4.20 Hasil Normalisasi Data Uji

USIA	IMT	GLUKOSA	RESISTIN
0,732	1	0,5	0
1	0,117	0,147	0,008
0	0	0	0,005
0,902	0,309	1	0,767
0,195	0,599	0,588	1
0,756	0,628	0,206	0,272

3. Set nilai $k = 1$.
 4. Menghitung nilai matriks H_Init dengan Persamaan 2.6.
- Berikut matriks H_Init yang didapatkan untuk $k=1$.

Tabel 4.21 Matriks H_Init pada $k = 1$

1,057	-0,085	0,728	-1,062	-1,484
0,514	0,462	0,304	-1,002	-0,844
0,356	0,808	0,123	0,078	0,213
0,345	0,42	-0,323	-1,602	-0,818
0,352	1,015	-0,291	-0,914	0,113
0,702	0,534	0,397	-1,014	-0,784

5. Menghitung matriks H sebagai keluaran dari *hidden layer* dengan menggunakan Persamaan 2.7.

Keluaran dari data ke-1 pada *hidden* neuron ke-1:

$$H_{11} = \frac{1}{1 + \exp(-1,057)} = 0,742$$

Keluaran dari data ke-1 pada *hidden* neuron ke-2 adalah:

$$H_{12} = \frac{1}{1 + \exp(-(-0,085))} = 0,479$$

Berikut matriks H yang merupakan keluaran dari *hidden layer*.

Tabel 4.22 Hasil Keluaran *Hidden Layer*

0,742	0,479	0,674	0,257	0,185
0,626	0,613	0,575	0,269	0,301
0,588	0,692	0,531	0,519	0,553
0,586	0,603	0,42	0,168	0,306
0,587	0,734	0,428	0,286	0,528
0,669	0,631	0,598	0,266	0,313

6. Menghitung nilai \hat{Y} dengan menggunakan Persamaan 2.10.

Berikut matriks \hat{Y} untuk $k = 1$.

Tabel 4.23 Hasil Klasifikasi Data Uji pada $k = 1$

Data Ke	Y1	Y2	Max	Kelas	Label Prediksi	Label Aktual
1	0,653	0,229	0,653	1	Healthy control	Healthy control
2	0,841	0,122	0,841	1	Healthy control	Healthy control
3	0,866	0,119	0,866	1	Healthy control	Healthy control
4	-0,079	1,121	1,121	2	Pasien	Pasien
5	0,092	0,944	0,944	2	Pasien	Pasien
6	0,736	0,263	0,736	1	Healthy control	Pasien

7. Set $k = k + 1$.
8. Ulangi langkah ke-4 sampai dengan ke-6 sampai nilai k terpenuhi.
9. Setelah dilakukan pengujian sebanyak $k = 3$ didapatkan hasil matriks \hat{Y} untuk tiap k lalu dilakukan *voting* dengan memilih hasil klasifikasi secara mayoritas.

Matriks \hat{Y} untuk $k = 1$.

Tabel 4.24 Hasil Klasifikasi Data Uji pada $k = 1$

Data Ke	Y1	Y2	Max	Kelas	Label Prediksi	Label Aktual
1	0,653	0,229	0,653	1	Healthy control	Healthy control

Tabel 4.24 Hasil Klasifikasi Data Uji pada $k = 1$ (lanjutan)

2	0,841	0,122	0,841	1	Healthy control	Healthy control
3	0,866	0,119	0,866	1	Healthy control	Healthy control
4	-0,079	1,121	1,121	2	Pasien	Pasien
5	0,092	0,944	0,944	2	Pasien	Pasien
6	0,736	0,263	0,736	1	Healthy control	Pasien

Matriks \hat{Y} untuk $k = 2$.

Tabel 4.25 Hasil Klasifikasi Data Uji pada $k = 2$

Data Ke	Y1	Y2	Max	Kelas	Label Prediksi	Label Aktual
1	0,852	0,15	0,852	1	Healthy control	Healthy control
2	0,697	0,304	0,697	1	Healthy control	Healthy control
3	0,623	0,374	0,623	1	Healthy control	Healthy control
4	-0,032	1,029	1,029	2	Pasien	Pasien
5	0,202	0,794	0,794	2	Pasien	Pasien
6	0,769	0,232	0,769	1	Healthy control	Pasien

Matriks \hat{Y} untuk $k = 3$.

Tabel 4.26 Hasil Klasifikasi Data Uji pada $k = 3$

Data Ke	Y1	Y2	Max	Kelas	Label Prediksi	Label Aktual
1	0,746	0,261	0,746	1	Healthy control	Healthy control
2	0,794	0,166	0,794	1	Healthy control	Healthy control
3	0,793	0,201	0,793	1	Healthy control	Healthy control
4	0,031	0,98	0,98	2	Pasien	Pasien

Tabel 4.26 Hasil Klasifikasi Data Uji pada $k = 3$ (lanjutan)

5	0,553	0,429	0,553	1	Healthy control	Pasien
6	0,674	0,329	0,674	1	Healthy control	Pasien

10. Hasil matriks \hat{Y} setelah dilakukan *voting*.

Matriks hasil klasifikasi dengan *voting*.

Tabel 4.27 Hasil Klasifikasi Data Uji dari *Voting* Tiga Prediksi

Data Ke	Kelas			Voting	Label Prediksi	Label Aktual
	K = 1	K = 2	K = 3			
1	1	1	1	1	Healthy control	Healthy control
2	1	1	1	1	Healthy control	Healthy control
3	1	1	1	1	Healthy control	Healthy control
4	2	2	2	2	Pasien	Pasien
5	2	2	1	2	Pasien	Pasien
6	1	1	1	1	Healthy control	Pasien

4.3.3 Perhitungan Proses Evaluasi algoritme V-ELM

Tabel 4.28 adalah proses perhitungan evaluasi menggunakan metode *confusion matrix*.

Tabel 4.28 Hasil Evaluasi *Confusion Matrix*

Prediksi	Aktual	
	P	HC
P	2	0
HC	1	3

Berdasarkan hasil Tabel 4.27 maka dapat ditentukan nilai untuk isi dari Tabel 4.28. Setelah mendapatkan nilai dari Tabel 4.28 maka dapat digunakan untuk mencari hasil *sensitivity*, *specificity* dan akurasi menggunakan Persamaan 2.11 sampai dengan Persamaan 2.13.

4.4 Perancangan Pengujian

4.4.1 Pengujian Banyak *Hidden* Neuron

Skenario terhadap pengujian *hidden* neuron dilakukan guna mengetahui banyaknya *hidden* neuron yang optimal pada metode V-ELM.

Tabel 4.29 Pengujian Jumlah *Hidden* Neuron

Jumlah <i>Hidden</i> Neuron	Akurasi					
	Percobaan ke- <i>i</i>					
	1	2	3	4	5	Rata-rata
5						
10						
15						
20						
25						
30						
35						
40						
45						
50						

4.4.2 Pengujian Nilai *k*

Skenario pengujian nilai *k* dilakukan untuk mengetahui nilai *k* yang optimal dalam melakukan perulangan pada metode V-ELM.

Tabel 4.30 Pengujian Nilai *k*

Nilai <i>k</i>	Akurasi					
	Percobaan ke- <i>i</i>					
	1	2	3	4	5	Rata-rata
3						
7						
11						
15						
19						
23						

Tabel 4.30 Pengujian Nilai k (lanjutan)

Nilai k	Akurasi					
	Percobaan ke- i					
	1	2	3	4	5	Rata-rata
27						
31						
35						
55						
99						
199						

4.4.3 Pengujian Fungsi Aktivasi

Skenario pengujian fungsi aktivasi dilakukan untuk mengetahui fungsi aktivasi yang dapat memberikan nilai akurasi yang optimal pada metode VELM.

Tabel 4.31 Pengujian Fungsi Aktivasi

Nilai k	Akurasi					
	Percobaan ke- i					
	1	2	3	4	5	Rata-rata
Sigmoid						
Fourier						
Hyperbolic Tangen						
Linear						

BAB 5 IMPLEMENTASI

5.1 Kode Sumber

5.1.1 Implementasi Normalisasi

Data latih dan uji diunggah ke sistem dalam format csv, kemudian dilakukan normalisasi untuk meminimalkan rentang nilai antar fitur. Implementasi mengenai kode sumber untuk normalisasi ditunjukkan pada Algoritme 1.

	Algoritme 1: Normalisasi
1	public double[][] normalisasi(double[][] fitur) {
2	for (int j = 0; j < fitur[0].length; j++) {
3	double min = Double.MAX_VALUE;
4	double max = Double.MIN_VALUE;
5	for (int i = 0; i < fitur.length; i++) {
6	//Min
7	if (fitur[i][j] < min) {
8	min = fitur[i][j];
9	}
10	if (fitur[i][j] > max) {
11	max = fitur[i][j];
12	}
13	}
14	for (int i = 0; i < fitur.length; i++) {
15	fitur[i][j] = (fitur[i][j] - min) / (max - min);
16	}
17	}
18	return fitur;
19	}

Baris 2-13 berfungsi untuk menemukan nilai minimal dan maksimal dari setiap fitur.

Baris 14-16 berfungsi untuk menghitung nilai hasil normalisasi tiap data.

Baris 18-19 berfungsi untuk mengembalikan nilai hasil data yang telah dinormalisasi ke dalam ke dalam array 'fitur'.

5.1.2 Implementasi V-ELM

5.1.2.1 Inisialisasi Bobot

Bobot merupakan salah satu masukan pada metode V-ELM yang didapatkan secara acak dari suatu rentang nilai. Algoritme 2 menunjukkan kode sumber dari inisialisasi bobot.

	Algoritme 2: Inisialisasi Bobot
1	public double[][] initBobot(){
2	double[][] bobot = new double[hidden][jumlah_fitur];
3	for (int i = 0; i < bobot.length; i++) {
4	for (int j = 0; j < bobot[0].length; j++) {
5	bobot[i][j] = randvalue(-0.5, 0.5);
6	}
7	}
8	return bobot;
8	}

Baris 2 berfungsi untuk membuat matriks bobot yang berisi nilai acak dengan dimensi, baris sebanyak jumlah *hidden* neuron dan kolom sebanyak jumlah fitur.

Baris 3-6 berfungsi untuk mencari nilai bobot secara acak, lalu disimpan pada matriks 'bobot'.

Baris 7 berfungsi untuk mengembalikan nilai hasil bobot yang telah diacak ke dalam *array* 'bobot'.

5.1.2.2 Menghitung Bias

Proses selanjutnya yang dilakukan setelah mencari bobot adalah mencari nilai bias secara acak. Implementasi kode sumber untuk mencari nilai bias ditunjukkan pada Algoritme 3.

	Algoritme 3: Inisialisasi Bias
1	public double[] bias(){
2	double[] bias = new double[hidden];
3	for (int i = 0; i < bias.length; i++) {
4	bias[i]= randvalue(0, 1);
5	}
6	return bias;
	}

Baris 2 berfungsi untuk membuat matriks bias dengan dimensi sejumlah *hidden neuron*.

Baris 3-4 berfungsi untuk mencari nilai bias secara acak .

Baris 5 berfungsi untuk mengembalikan nilai bias yang telah diacak ke dalam *array* 'bias'.

5.1.2.3 Menghitung Keluaran Matriks *H*

Setelah mendapatkan bobot dan bias, proses selanjutnya adalah mencari nilai *H*. Implementasi kode sumber untuk menghitung nilai *H* ditampilkan pada Algoritme 4.

	Algoritme 4: Hitung Matriks Keluaran Hidden Layer
1	public double[][] h_init(double[][] fitur, double[][]
2	bobot, double[] bias){
3	Matrix w = new Matrix(bobot);
4	Matrix w_trans = w.transpose();
5	Matrix x = new Matrix(fitur);
6	Matrix x_wt = x.times(w_trans);
7	double[][] h_init = x_wt.toArray();
8	for (int i = 0; i < h_init.length; i++) {
9	for (int j = 0; j < h_init[0].length; j++) {
10	h_init[i][j] += bias[j];
11	}
12	}
13	double[][] h = new double[h_init.length][h_init[0].length];
14	for (int i = 0; i < h.length; i++) {
15	for (int j = 0; j < h[0].length; j++) {
16	h[i][j] = 1/(1+Math.exp(-h_init[i][j]));
17	}
18	}
19	return h;
20	}

Baris 3-5 berfungsi untuk menyimpan bobot, bobot *transpose*, data ke dalam bentuk Matrix JAMA.

Baris 6 berfungsi untuk menghitung hasil perkalian antara nilai fitur dengan bobot *transpose*.

Baris 7 berfungsi untuk menyimpan hasil dari baris 6 ke dalam matriks *h_init*.

Baris 8-12 berfungsi untuk melakukan operasi penjumlahan antara matriks *h_init* dengan bias.

Baris 14-18 berfungsi untuk menghitung matriks *h_init* dengan fungsi sigmoid biner lalu disimpan ke dalam matriks *h*.

Baris 19 berfungsi untuk mengembalikan nilai matriks *H* ke dalam array 'h'.

5.1.2.4 Menghitung Moore Penrose Matriks H

Setelah mendapatkan nilai matriks *h*, proses selanjutnya adalah mencari nilai H^+ . Implementasi dari proses menghitung nilai H^+ ditunjukkan pada Algoritme 5.

	Algoritme 5: Moore Penrose
1	public Matrix moore_penrose(double[][] h) {
2	Matrix H = new Matrix(h);
3	Matrix HT = H.transpose();
4	Matrix HT_H = HT.times(H);
5	Matrix in_HT_H = HT_H.inverse();
6	Matrix H_plus = in_HT_H.times(HT);
7	return H_plus;
8	}

Baris 2 berfungsi sebagai perintah untuk menyimpan matriks *H* ke dalam Matrix JAMA.

Baris 3 berfungsi sebagai perintah untuk menyimpan hasil transpose dari Matrix *H*.

Baris 4-6 berfungsi sebagai perintah untuk mencari nilai H^+ .

Baris 7 berfungsi untuk mengembalikan nilai H^+ yang telah dihitung ke dalam array 'H_plus'.

5.1.2.5 Menghitung Beta

Tahapan yang dilakukan setelah mendapatkan nilai matriks H^+ adalah menghitung nilai beta. Implementasi ke dalam kode sumber dari proses menghitung beta ditampilkan pada Algoritme 6.

	Algoritme 6: Hitung Output Weight
1	public Matrix beta_topi(Matrix H_plus, double[][] target) {
2	Matrix Target = new Matrix(target);
3	Matrix beta = H_plus.times(Target);
4	return beta;
5	}

Baris 2 berfungsi untuk menyimpan matriks target ke dalam Matrix JAMA.

Baris 3 berfungsi untuk menghitung nilai beta dengan melakukan perkalian antara Matrix *H_plus* dengan Matrix target.

Baris 4 berfungsi untuk mengembalikan nilai beta yang telah dihitung ke dalam array 'beta'.

5.1.2.6 Menghitung Keluaran *Output Layer*

Untuk mendapatkan nilai keluaran *output layer* diperlukan masukan berupa matriks H dan nilai beta. Implementasi dari proses untuk mendapatkan keluaran *output layer* ke dalam kode sumber ditampilkan pada Algoritme 7.

	Algoritme 7: Hitung Keluaran Output Layer
1	public String[] output_layer(double[][] h, Matrix beta) {
2	Matrix H = new Matrix(h);
3	Matrix Y = H.times(beta);
4	double[][] y = Y.getArray();
5	String[] kelas = new String[y.length];
6	for (int i = 0; i < y.length; i++) {
7	if (y[i][0] > y[i][1]) {
8	kelas[i] = "HC";
9	} else {
10	kelas[i] = "P";
11	}
12	}
13	return kelas;
14	}

Baris 2 berfungsi untuk menyimpan matriks H ke dalam Matrix JAMA.

Baris 3 berfungsi untuk melakukan perhitungan hasil klasifikasi dengan melakukan perkalian antara Matrix H dengan Matrix beta.

Baris 6-12 berfungsi untuk memberi label kelas sebagai hasil klasifikasi.

5.1.2.1 Menghitung Keluaran *Voting*

Dalam melakukan proses klasifikasi, metode V-ELM akan dilakukan sampai dengan batas nilai k terpenuhi. Implementasi kode sumber untuk mendapatkan keluaran *voting* ditunjukkan pada Algoritme 8.

	Algoritme 8: Voting Hasil Klasifikasi
1	public String[] vote(ArrayList<String[]> hasil, double[][]
2	fitur){
3	String[][] kelas = new String[hasil.size()][fitur.length];
4	String[] label = new String[fitur.length];
5	for (int i = 0; i < kelas.length; i++) {
6	kelas[i] = hasil.get(i);
7	}
8	for (int j = 0; j < kelas[0].length; j++) {
9	int hc=0, p=0;
10	//votingan tiap k
11	for (int i = 0; i < kelas.length; i++) {
12	if (kelas[i][j].equalsIgnoreCase("HC")) {
13	hc++;
14	} else if (kelas[i][j].equalsIgnoreCase("P")) {
15	p++;
16	}
17	}
18	if (hc > p) {
19	label[j] ="HC";
20	} else if (hc < p) {

```

21         label[j] ="P";
22     }
23 }
24     return label;
25 }

```

Baris 5-7 berfungsi untuk mendapatkan banyak data yang akan dilakukan proses *voting*.

Baris 11-17 berfungsi untuk menghitung hasil *voting* kelas setiap data .

Baris 18-22 berfungsi untuk menentukan kelas untuk setiap data.

5.2 Antarmuka

Agar hasil keluaran dari sistem klasifikasi dapat diketahui, maka hasil klasifikasi akan ditampilkan seperti pada Gambar 5.1.

Data	Prediksi	Aktual
1	HC	HC
2	HC	HC
3	HC	HC
4	HC	HC
5	P	HC
6	HC	HC
7	HC	P
8	P	P
9	P	P
10	P	P
11	P	P
12	P	P

Gambar 5.1 Antarmuka Hasil Klasifikasi

Untuk mengetahui hasil evaluasi dari klasifikasi menggunakan V-ELM maka ditampilkan seperti Gambar 5.2.

Confussion Matrix

TP = 12.0	FP = 2.0
FN = 1.0	TN = 8.0

Hasil Evaluasi

Sensitivity	= 92,31
Akurasi	= 86,96
Specificity	= 80,00

Gambar 5.2 Antarmuka Hasil Evaluasi

BAB 6 PENGUJIAN DAN ANALISIS

6.1 Pengujian

Berdasarkan perancangan pengujian yang telah ditentukan pada Bab 4, berikut adalah daftar pengujian yang akan dilakukan pada penelitian ini.

1. Pengujian Banyak *Hidden Neuron*
2. Pengujian Fungsi Aktivasi
3. Pengujian Nilai k

Setiap skenario pengujian yang dilakukan akan diulangi sebanyak 5 kali percobaan. Pada ketiga pengujian tersebut, perbandingan banyak data latih dan data uji adalah sebesar 80:20, karena ELM merupakan jaringan yang memiliki karakteristik jaringan pembelajaran, sehingga semakin banyak pola data yang dipelajari maka akan semakin baik hasil yang diberikan (Sugianto et al., 2018). Setelah setiap pengujian dilakukan maka akan diberikan analisis berdasarkan dari hasil pengujian.

6.2 Hasil dan Analisis Pengujian

6.2.1 Pengujian Banyak *Hidden Neuron*

Pengujian banyak *hidden neuron* dilakukan untuk mengetahui pengaruh banyaknya *hidden neuron* terhadap hasil proses klasifikasi.

Nilai parameter lainnya yang digunakan sebagai berikut.

1. Nilai k : 3
2. Fungsi aktivasi : *Sigmoid*
3. Rasio data latih dan data uji : 80% : 20%

Tabel 6.1 Hasil Pengujian Banyak *Hidden Neuron*

Jumlah <i>Hidden Neuron</i>	Akurasi					Rata-rata
	Percobaan ke- <i>i</i>					
	1	2	3	4	5	
5	56,52	69,57	69,57	56,52	73,91	65,218
10	69,57	82,61	78,26	86,96	86,96	80,872
15	73,91	86,96	82,61	95,65	78,26	83,478
20	82,61	86,96	91,30	82,61	78,26	84,348
25	73,91	95,65	82,61	73,91	82,61	81,738
30	82,61	69,57	73,91	86,96	73,91	77,392
35	69,57	73,91	82,61	82,61	82,61	78,262

Tabel 6.1 Hasil Pengujian Banyak *Hidden* Neuron (lanjutan)

Jumlah <i>Hidden</i> Neuron	Akurasi					Rata-rata
	Percobaan ke- <i>i</i>					
	1	2	3	4	5	
40	78,26	56,52	73,91	73,91	82,61	73,042
45	69,57	69,57	69,57	82,61	56,52	69,568
50	69,57	65,22	47,83	69,57	65,22	63,482

Hasil pengujian banyak *hidden* neuron pada Tabel 6.1 ditampilkan dalam representasi grafik seperti pada Gambar 6.1.



Gambar 6.1 Grafik Hasil Pengujian Banyak *Hidden* Neuron

Pada Gambar 6.1 menunjukkan bahwa banyak *hidden* neuron pada rentang nilai 5 sampai dengan 20 menghasilkan nilai akurasi yang meningkat. Penurunan hasil nilai akurasi terjadi pada *hidden* neuron dengan rentang nilai 21 sampai 30, lalu hasil nilai akurasi kembali mengalami peningkatan pada *hidden* neuron dengan rentang nilai 31 sampai 35, hal tersebut dapat terjadi karena pada metode V-ELM, inisialisasi parameter masukan berupa bobot dan bias dilakukan secara acak. Pertambahan banyak *hidden* neuron mengakibatkan bertambahnya pola yang terbentuk pada tahap pembelajaran sehingga memungkinkan terjadi *overfitting*. Begitupula jika banyaknya *hidden* neuron terlalu sedikit, maka pola yang dikenali akan terbatas sehingga terjadi *underfitting*. Baik *overfitting* maupun *underfitting* akan menyebabkan hasil akurasi cenderung rendah.

6.2.2 Pengujian Fungsi Aktivasi

Pengujian mengenai fungsi aktivasi dilakukan guna mengetahui pengaruh penggunaan fungsi aktivasi yang memberikan hasil tingkat akurasi optimal dalam melakukan klasifikasi kanker payudara. Sedangkan untuk parameter masukan

seperti banyak *hidden* neuron menggunakan hasil pengujian banyak *hidden* neuron.

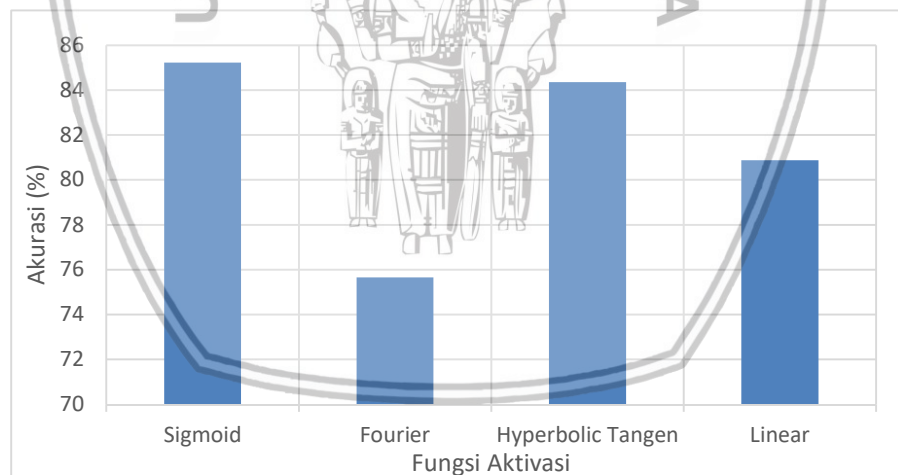
Nilai parameter lainnya yang digunakan sebagai berikut.

1. Banyak *hidden* neuron : 20
2. Nilai k : 3
3. Rasio data latih dan data uji : 80% : 20%

Tabel 6.2 Pengujian Fungsi Aktivasi

Fungsi Aktivasi	Akurasi					Rata-rata Akurasi
	Percobaan ke- <i>i</i>					
	1	2	3	4	5	
Sigmoid	82,61	82,61	91,30	82,61	86,96	85,218
Fourier	65,22	78,26	65,22	86,96	82,61	75,654
Hyperbolic Tangen	82,61	82,61	86,96	86,96	82,61	84.35
Linear	82,61	86,96	78,26	73,91	82,61	80,87

Hasil pengujian fungsi aktivasi pada Tabel 6.2 ditampilkan dalam representasi grafik seperti pada Gambar 6.2.



Gambar 6.2 Grafik Hasil Pengujian Fungsi Aktivasi

Pada Gambar 6.2 menampilkan hasil rata-rata tingkat akurasi yang didapat dari empat jenis pengujian fungsi aktivasi. Penggunaan fungsi aktivasi *sigmoid* dapat menghasilkan tingkat akurasi yang paling baik dibanding fungsi aktivasi lainnya. Rata-rata tingkat akurasi yang dicapai dari fungsi aktivasi *sigmoid* adalah 85,218%. Hal tersebut menunjukkan bahwa fungsi aktivasi *sigmoid* sesuai untuk digunakan pada metode dan data dalam penelitian ini. Hasil akurasi yang baik didapatkan ketika fungsi aktivasi yang digunakan fungsi aktivasi *sigmoid* dan *hyperbolic tangen* karena keduanya bagian dari fungsi *logistic* (Shenouda, 2006).

6.2.3 Pengujian Nilai k

Pengujian mengenai nilai k yang merepresentasikan banyaknya perulangan metode ELM dalam V-ELM dilakukan untuk mengetahui pengaruh nilai k terhadap hasil akurasi dari proses klasifikasi. Sedangkan parameter masukan *hidden* neuron yang digunakan pada skenario percobaan ini berdasarkan hasil akurasi terbaik yang diperoleh dari pengujian banyak *hidden neuron* dan pengujian fungsi aktivasi sebelumnya.

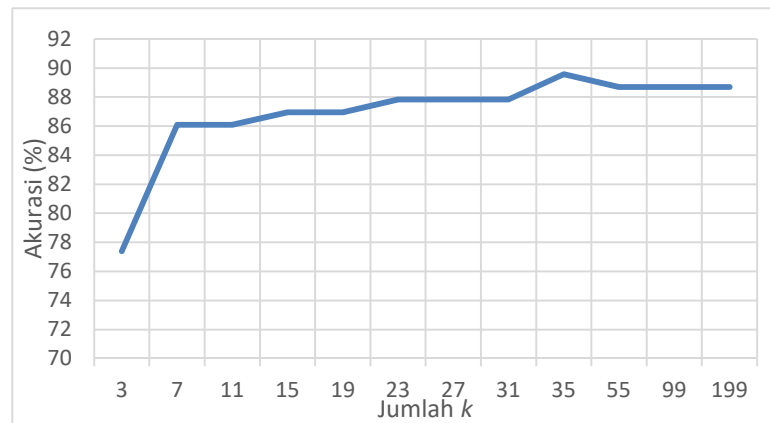
Nilai parameter lainnya yang digunakan sebagai berikut:

1. Banyak *hidden* neuron : 20
2. Fungsi Aktivasi : *Sigmoid*
3. Perbandingan data latih dan data uji : 80% : 20%

Tabel 6.3 Hasil Pengujian Nilai k

Jumlah Nilai k	Akurasi					Rata-rata
	Percobaan ke- i					
	1	2	3	4	5	
3	73,91	78,26	82,61	73,91	78,26	77,39
7	82,61	91,3	86,96	86,96	82,61	86,088
11	82,61	86,96	86,96	86,96	86,96	86,09
15	91,3	86,96	91,3	82,61	82,61	86,956
19	91,3	82,61	86,96	86,96	86,96	86,958
23	86,96	86,96	91,3	86,96	86,96	87,828
27	82,61	91,3	91,3	86,96	86,96	87,826
31	86,96	91,3	95,65	78,26	86,96	87,826
35	86,96	91,3	91,3	86,96	91,3	89,564
55	91,3	86,96	86,96	86,96	91,3	88,696
99	86,96	91,3	86,96	86,96	91,3	88,696
199	86,96	86,96	91,3	86,96	91,3	88,696

Hasil pengujian nilai k pada Tabel 6.3 ditampilkan dalam representasi grafik seperti pada Gambar 6.3.



Gambar 6.3 Grafik Hasil Pengujian Nilai k

Gambar 6.3 menampilkan tingkat akurasi yang diperoleh dari pengujian nilai k . Dari lima kali percobaan yang telah dilakukan, untuk setiap nilai k diperoleh bahwa bertambah besarnya nilai k yang digunakan dapat menghasilkan rata-rata tingkat akurasi yang tinggi karena semakin banyak proses *voting* yang dilakukan memilih label kelas yang dominan maka sistem dapat mengklasifikasikan dengan benar sehingga rata-rata tingkat akurasi yang akan diperoleh dapat lebih tinggi dibanding nilai k yang kecil. Dari pengujian yang telah dilakukan 35 merupakan nilai k yang paling optimal dalam penelitian ini dengan menghasilkan rata-rata tingkat akurasi sebesar 89,564%.

6.2.4 Hasil Pengujian dan Analisis Global

Setelah melakukan tiga skenario percobaan, dapat diketahui bahwa penggunaan banyak *hidden* neuron sangat menentukan tingkat akurasi yang akan dicapai oleh sistem. Nilai akurasi akan rendah bila menggunakan *hidden* neuron yang terlalu sedikit ataupun banyak. Hal tersebut dapat dikaitkan dengan gejala *underfitting* dan *overfitting*, apabila banyak *hidden* neuron yang digunakan terlalu sedikit maka sistem tidak dapat mengenali pola data pada tahap pelatihan sehingga pada saat dilakukan pengujian akan menghasilkan akurasi yang rendah sehingga terjadi *underfitting*, lalu jika banyak *hidden* neuron yang digunakan terlalu banyak maka sistem akan mengenali pola data pada tahap pelatihan terlalu banyak sehingga pada saat dilakukan pengujian akan menghasilkan akurasi yang rendah sehingga terjadi *overfitting*. Pemilihan nilai k dengan bilangan ganjil yang semakin besar dapat meningkatkan rata-rata tingkat akurasi. Hal tersebut dikarenakan semakin banyaknya proses *voting* dilakukan untuk memilih label yang dominan maka peluang sistem dalam mengklasifikasikan akan semakin optimal. Penggunaan fungsi aktivasi juga berpengaruh terhadap hasil akurasi yang optimal, yang mana fungsi sigmoid biner merupakan fungsi aktivasi yang akan menghasilkan nilai akurasi yang optimal dibanding menggunakan fungsi aktivasi lainnya. Dengan menggunakan parameter optimal dari ketiga pengujian sebelumnya, yaitu 20 *hidden* neuron, fungsi aktivasi Sigmoid dan nilai $k=35$, didapatkan hasil akurasi sebesar 89,564%, *sensitivity* sebesar 96,924%, *specificity* sebesar 80%.

BAB 7 PENUTUP

7.1 Kesimpulan

Beberapa hal yang dapat disimpulkan setelah melakukan tahap perancangan, implementasi, pengujian dan analisis adalah sebagai berikut.

1. Metode V-ELM telah dirancang dan dapat diimplementasikan pada permasalahan klasifikasi pengidap kanker payudara. Proses pertama yang dilakukan adalah membagi data latih dan data uji. Kemudian melakukan proses normalisasi terhadap data latih dan data uji. Setelah itu dilakukan klasifikasi dengan metode V-ELM, sehingga didapatkan hasil klasifikasi dari data uji.
2. Dari hasil pengujian didapatkan nilai optimal untuk banyak *hidden* neuron adalah 20 neuron, nilai optimal untuk k adalah 35 dan fungsi aktivasi yang memberikan hasil paling optimal adalah fungsi aktivasi *Sigmoid*.
3. Dengan menggunakan parameter optimal dari hasil pengujian, didapatkan rata-rata akurasi sebesar 89,564% serta hasil evaluasi lainnya seperti rata-rata *sensitivity* mencapai 96,924%, rata-rata *specificity* sebesar 80%. Hasil akurasi yang tinggi ini mengindikasikan bahwa metode V-ELM cocok digunakan untuk klasifikasi data pengidap kanker payudara *Coimbra dataset* 2018.

7.2 Saran

Penulis memberikan saran yang dapat dilakukan untuk kelanjutan penelitian ini, antara lain:

1. Untuk meningkatkan rata-rata nilai akurasi diharapkan banyak data yang digunakan ditambah dan banyak data pada tiap kelas seimbang, karena dalam penelitian ini banyak data yang digunakan sebanyak 116 data yang mana untuk label kelas pasien sebanyak 64 data dan label kelas *healthy control* sebanyak 52 data.
2. Pengujian fungsi aktivasi pada algoritme V-ELM dapat menggunakan beberapa jenis fungsi aktivasi secara simultan.

DAFTAR REFERENSI

- Adinugroho, S., Sari, Y. A., 2018. *Implementasi Data Mining Menggunakan Weka*. Malang: UB Press.
- Alodokter, 2018. *Mengenal Anatomi Payudara Wanita*. [online] Tersedia di:< <https://www.alodokter.com/mengenal-anatomi-payudara-wanita>> [Diakses 25 Agustus 2018]
- Alodokter, 2014. *Kanker payudara*. [online] Tersedia di:< <https://www.alodokter.com/kanker-payudara>> [Diakses 25 Agustus 2018]
- Cao, J., Lin, Z., Huang, G. & Liu, N., 2012. Voting based extreme learning machine. *Information Sciences*, 185, p.66-77.
- Daly, C., Rollins, B. J., 2003. Monocyte Chemoattractant Protein-1 (CCL2) in Inflammatory Disease and Adaptive Immunity: Therapeutic Opportunities and Controversies. *Microcirculation*, 10, p.247-257.
- Ding, S., Zhao, H., Zhang, Y. & Nie, R., 2015. *Extreme Learning Machine: Algorithm, Theory and Applications*. *Artificial Intelligence*, 44, p.103-115.
- Dewi, S. N., Cholissodin, I., & Santoso, E., 2018. Prediksi Jumlah Kriminalitas Menggunakan Metode Extreme Learning Machine (Studi Kasus Di Kabupaten Probolinggo). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(11), p.4687-4693.
- Fadilla, I., Adikara, P. P., & Perdana, R. S., 2018. Klasifikasi Penyakit Chronic Kidney Disease (CKD) Dengan Menggunakan Metode Extreme Learning Machine (ELM). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(10), p.3397-3405.
- Ginting, S. E. B., Widodo, A. W., & Adikara, P. P., 2018. Voting Based Extreme Learning Machine dalam Klasifikasi Computer Network. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(6), p.2158-2167.
- Huang, G., Huang, G.-B., Song, S., & You, K., 2015. *Trends in Extreme Learning Machines: A Review*. *Neural Networks*, 61, p.32-48.
- Joyce, L. K., 2006. *Buku Saku Pemeriksaan Laboratorium dan Diagnostic dengan Implikasi Keperawatan*. Jakarta: EGC.
- Junaidi, I., 2014. *Hidup Sehat Bebas Kanker – Mewaspada Kanker Sejak dari Dini*. Yogyakarta: Rapha Publishing.
- Liputan6, 2016. *Lovepink dan Pinkvoice Ajak Masyarakat Bicarakan Bahaya Kanker Payudara*. [online] Tersedia di:< <https://www.liputan6.com/lifestyle/read/3658704/lovepink-dan-pinkvoice-ajak-masyarakat-bicarakan-bahaya-kanker-payudara>> [Diakses 15 Oktober 2018]

- Matin, S. S. dan Veria, V. A., 2013. *Body Mass Index (BMI) Sebagai Salah Satu Faktor yang Berkontribusi Terhadap Prestasi Belajar Remaja*. Jurnal VISIKES, 12(2), p.163-169.
- Patricio, M., Pereira, J., Crisóstomo, J., Metafome, P., Gomes, M., Seica, R., & Caramelo, F., 2018. *Using Resistin, glucose, age and BMI to predict the presence of breast cancer*. *Informaton Sciences*, 185, 66-77.
- Pusat Data dan Informasi Kementrian Kesehatan Republik Indonesia, 2016. *Infodatin: Bulan Peduli Kanker Payudara*. [pdf] Kementerian Kesehatan . Tersedia di: <
http://www.depkes.go.id/download.php?file=download/pusdatin/infodatin/InfoDatin%20Bulan%20Peduli%20Kanker%20Payudara_2016.pdf> [Diakses 21 Agustus 2018]
- Pratama, M. I., 2018. Peramalan Harga Saham Menggunakan Metode Extreme Learning Machine (ELM) Studi Kasus Saham Bank Mandiri. S1. Universitas Brawijaya.
- Qiu, S., Gao, L., Wang, J., 2015. Classification and Regression of ELM, LVQ and SVM for E-nose Data of Strawberry Juice. *Journal of Food Engineering*, 144, p.77-85.
- Shenouda, E., 2006. A Quantitative Comparison of Different MLP Activation Functions in Classification. *Advances in Neural Networks – ISNN 2006*, 3971, p.849-857.
- Shuldiner, A. R., Yang, R., Gong, D., 2001. *Resistin, Obesity and Insulin Resistance – The Emerging Role of The Adipocyte as an Endocrine Organ*. *N Engl J Med*, 345(18), p.1345-1346.
- Sugianto, N. A., Cholissodin, I., & Widodo, A. W., 2018. Klasifikasi Keminatan Menggunakan Algoritme Extreme Learning Machine dan Particle Swarm Optimization untuk Seleksi Fitur. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(5), p.1856-1865.
- Suyanto, 2014. *Artificial Intelligence: Searching, Reasoning, Planning and Learning*. Bandung: Informatika.
- Wang, H., Zheng, B., Yoon, S. W., & Ko, H. S., 2018. A Support Vector Machine-Based Ensemble Algorithm For Breast Cancer Diagnosis. *European Journal of Operational Research*, 267, p.687-699.
- WHO, 2018. Breast Cancer. [online] Tersedia di:<
<http://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/>> [Diakses 15 Oktober 2018]
- Wicaksono, Y. H., 2018. Identifikasi Awal Pengguna Narkoba Menggunakan Metode Learning Vector Quantization (LVQ). S1. Universitas Brawijaya.
- Wolipop, 2012. *8 Hormon Paling Berpengaruh yang Menentukan Berat Badan Anda*. [online] Tersedia di: <<https://wolipop.detik.com/health-and-diet/d->

1953450/8-hormon-paling-berpengaruh-yang-menentukan-berat-badan-anda> [Diakses 25 Oktober 2018]

